

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 1, 2003, 16:02:29; Search time 25 seconds

(without alignments)
3262.219 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLETNPLSPETAMREPAP.....VOKKGRYYELVQSLGKH 1334

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 248812 seqs, 61136040 residues

Total number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Published Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6721	100.0	1334	9	US-09-758-828-2
2	2489	37.0	1280	9	US-10-072-621-7
3	2489	37.0	1280	10	US-09-866-866A-2
4	2487.5	37.0	1272	10	US-09-769-097-4
5	2482	36.9	1280	9	US-10-044-671-2
6	2480	36.9	1280	10	US-09-866-866A-4
7	2475.5	36.8	1272	10	US-09-769-097-2
8	2475.5	36.8	1276	10	US-09-866-866A-8
9	2453.5	36.5	1276	10	US-09-866-866A-6
10	2282	34.0	1195	10	US-09-873-409-6
11	2258.5	33.6	1222	10	US-09-873-409-5
12	2160.5	32.1	1275	10	US-09-749-340-6
13	2067	30.8	1058	9	US-09-873-409-4
14	1940	28.9	1263	9	US-09-882-694-11
15	1690.5	25.2	812	10	US-09-873-409-2
16	1617	24.1	1394	9	US-10-101-388-3
17	1254	18.7	659	10	US-09-873-409-1
18	1001.5	14.9	1247	9	US-09-738-626-4751
19	937	13.9	514	10	US-09-873-409-8

20	913.5	13.6	541	10	US-09-873-409-7	Sequence 7, Appl1
21	814.5	12.1	766	9	US-10-072-621-6	Sequence 6, Appl1
22	813.5	12.1	766	9	US-10-156-239-17	Sequence 17, Appl1
23	813.5	12.1	766	10	US-09-795-693-17	Sequence 17, Appl1
24	756	11.2	582	10	US-09-815-242-14006	Sequence 14006, A
25	749.5	11.2	582	10	US-09-815-242-10115	Sequence 10115, A
26	740.5	11.0	1464	9	US-10-012-896-1008	Sequence 1008, Ap
27	739.5	11.0	1464	9	US-10-012-896-1009	Sequence 538, App
28	739.5	11.0	1461	9	US-10-012-896-1009	Sequence 1009, Ap
29	739.5	11.0	1261	9	US-09-895-793-538	Sequence 538, App
30	739.5	11.0	1261	9	US-09-895-814-538	Sequence 538, App
31	739.5	11.0	1261	10	US-09-753-143-538	Sequence 538, App
32	739.5	11.0	1261	10	US-09-780-669-538	Sequence 538, App
33	739.5	11.0	1361	10	US-09-822-827-538	Sequence 538, App
34	720	10.7	1361	10	US-10-087-782A-31	Sequence 31, Appl1
35	702	10.4	1360	9	US-09-934-421A-2	Sequence 2, Appl1
36	702	10.4	1360	9	US-10-162-012-34	Sequence 34, Appl1
37	691.5	10.3	603	10	US-09-815-242-5202	Sequence 5202, Ap
38	682	10.1	748	9	US-09-870-759-41	Sequence 41, Appl1
39	674.5	10.0	1437	9	US-10-154-452-2	Sequence 2, Appl1
40	674.5	10.0	1437	9	US-09-934-421A-6	Sequence 6, Appl1
41	674.5	10.0	1437	9	US-10-162-012-38	Sequence 38, Appl1
42	670	10.0	1228	9	US-10-012-896-537	Sequence 537, App
43	670	10.0	1228	9	US-09-895-793-537	Sequence 537, App
44	670	10.0	1228	9	US-09-895-814-537	Sequence 537, App
45	670	10.0	1228	10	US-09-759-143-537	Sequence 537, App

ALIGNMENTS

RESULT 1
US-09-758-828-2
Sequence 2, Application US/09758828
Patent No. US20020164772A1
GENERAL INFORMATION:
APPLICANT: Skatrud, Paul L.
de Waard, Maarten A.
Peery, Robert B.
Andrade, Alan C.
TITLE OF INVENTION: Multiple Drug Resistance Gene atrod of
Aspergillus nidulans
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Eli Lilly and Company
STREET: Lilly Corporate Center
CITY: Indianapolis
STATE: Indiana
COUNTRY: U.S.
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/758, 828
FILING DATE: 11-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,545
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Webster, Thomas D.
REGISTRATION NUMBER: 39,872
REFERENCE/DOCKET NUMBER: X-11766
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3334
TELEFAX: 317-276-2763
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1334 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-758-828-2

Query Match 100.0%; Score 6721; DB 9; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSPLTNPLSPETAMREPASTSTTEQASTPHADEKTLISLSPASSSTTAPADKENRP 60
DB 1 MSPLTNPLSPETAMREPASTSTTEQASTPHADEKTLISLSPASSSTTAPADKENRP 60
QY 61 KSSSSNNVAVNEVDALIAHPEDEKQVLTQLEIKVNIISFGIMARVATKADILIMVS 120
DB 61 KSSSSNNVAVNEVDALIAHPEDEKQVLTQLEIKVNIISFGIMARVATKADILIMVS 120
QY 121 TICATAAASFQRMILYOISYDEFDLKNLYEYVIGIGEVTVYVSTVGFIYGEHA 180
DB 121 TICATAAASFQRMILYOISYDEFDLKNLYEYVIGIGEVTVYVSTVGFIYGEHA 180
QY 181 TOKIREYVLESILRONIGYFDKLAGEVYTRITADTNLIQDISEKVGILTALATFYTA 240
DB 181 TOKIREYVLESILRONIGYFDKLAGEVYTRITADTNLIQDISEKVGILTALATFYTA 240
QY 241 FIATVYKWKIALICSTTVALTVMGGSOPIIKYKSKLDSYGAGTVAEEVISITRN 300
DB 241 FIATVYKWKIALICSTTVALTVMGGSOPIIKYKSKLDSYGAGTVAEEVISITRN 300
QY 301 ATAFGTQDKLAKQYEVHIDEAEKMGTKNOIYVGFIMIGAMFGLMNSYGLGFWMGSRFLVD 360
DB 301 ATAFGTQDKLAKQYEVHIDEAEKMGTKNOIYVGFIMIGAMFGLMNSYGLGFWMGSRFLVD 360
QY 361 GAVDYGDLITVYMAILLISFSLGNVSPNAQATNAVAAAATFGITIDROSPIDPYSNECK 420
DB 361 GAVDYGDLITVYMAILLISFSLGNVSPNAQATNAVAAAATFGITIDROSPIDPYSNECK 420
QY 421 TLDFEGHIELKNVHIYPSREPTVMEVSLSPAGKTTALVSGSGSKSVYGVYERF 480
DB 421 TLDFEGHIELKNVHIYPSREPTVMEVSLSPAGKTTALVSGSGSKSVYGVYERF 480
QY 481 YMPVGTVLDDHDIKIDNLRLROOISLVSOEPLYFGTTIKNIRHGLIGTKYENESD 540
DB 481 YMPVGTVLDDHDIKIDNLRLROOISLVSOEPLYFGTTIKNIRHGLIGTKYENESD 540
QY 541 KYRELIEENAKKANAHDFTALPEGETNVGQRFLLSGGQKORIALAARAVSDPKILL 600
DB 541 KYRELIEENAKKANAHDFTALPEGETNVGQRFLLSGGQKORIALAARAVSDPKILL 600
QY 601 DEATSALDTKSGVVOALERAAGRTTIVIAHRLSTIKTANIVVVGKTAEOGTDE 660
DB 601 DEATSALDTKSGVVOALERAAGRTTIVIAHRLSTIKTANIVVVGKTAEOGTDE 660
QY 661 LVDRGAVKRLVEAQRINEQKADALADADADLTNADIAKIKITASASDLDGKPTTID 720
DB 661 LVDRGAVKRLVEAQRINEQKADALADADADLTNADIAKIKITASASDLDGKPTTID 720
QY 721 RTGTHKSVSAILSKRPETTPKYSMLTLKFVASFNREPIYMLIGLVFSLAGGGOPT 780
DB 721 RTGTHKSVSAILSKRPETTPKYSMLTLKFVASFNREPIYMLIGLVFSLAGGGOPT 780
QY 781 QAVYAKAISTLSPESQSKLRHADFMSLMEFVVGITOTOSTNGAFAVCSERLIR 840
DB 781 QAVYAKAISTLSPESQSKLRHADFMSLMEFVVGITOTOSTNGAFAVCSERLIR 840
QY 841 RARSTAFPTILRODIAFDEKENSSTGALTSLSTETKHLSGVSGTTLGILMTSTLGA 900
DB 841 RARSTAFPTILRODIAFDEKENSSTGALTSLSTETKHLSGVSGTTLGILMTSTLGA 900
QY 901 IIALAIGWKIALVCISSVPLVLAAGFYRFYLAQFQSRSKLAYSANFACEATSSIRT 960
DB 901 IIALAIGWKIALVCISSVPLVLAAGFYRFYLAQFQSRSKLAYSANFACEATSSIRT 960

QY 961 VASLTRERDWEIYHAQIDAOGRSTLSYLSBSLLYASSQAALVFECVALGFVGGTLLGH 1020
DB 961 VASLTRERDWEIYHAQIDAOGRSTLSYLSBSLLYASSQAALVFECVALGFVGGTLLGH 1020
QY 1021 HEYDIFREFVCESELLFGAOSAGTFFSFAPDMGKAKNAAEFRRLFDKRPQIDNNSEGE 1080
DB 1021 HEYDIFREFVCESELLFGAOSAGTFFSFAPDMGKAKNAAEFRRLFDKRPQIDNNSEGE 1080
QY 1081 KLEIVGEIEFRNVHFRPTREPQVLRGLDITVRPGQYVALVPSGCKSTTIALERE 1140
DB 1081 KLEIVGEIEFRNVHFRPTREPQVLRGLDITVRPGQYVALVPSGCKSTTIALERE 1140
QY 1141 YDAINGSLVDGKDISKININSRPSLSVSOEPLTYOGTIENTILLGIVEDVBEELI 1200
DB 1141 YDAINGSLVDGKDISKININSRPSLSVSOEPLTYOGTIENTILLGIVEDVBEELI 1200
QY 1201 KACKDANIYDFIMSLPEGEFTVVGSKGMLSGGQKORVIAARALLRDPKILLDATSAL 1260
DB 1201 KACKDANIYDFIMSLPEGEFTVVGSKGMLSGGQKORVIAARALLRDPKILLDATSAL 1260
QY 1261 DSESEKVOALDAARAGRTTIAVAHRLSTIQKADVIYFDGKIVESGTHSELVQKGR 1320
DB 1261 DSESEKVOALDAARAGRTTIAVAHRLSTIQKADVIYFDGKIVESGTHSELVQKGR 1320
QY 1321 YVELVNIOSLGKH 1334
DB 1321 YVELVNIOSLGKH 1334

RESULT 2
US-10-072-621-7
Sequence 7, Application US/10072621
Patent No. US20020169137A1
GENERAL INFORMATION:
APPLICANT: Reiner, Peter B.
APPLICANT: Connop, Bruce P.
APPLICANT: Pollard, Michelle
TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
FILE REFERENCE: 100103.402
CURRENT FILING DATE: 2002-02-08
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 1280
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-621-7

Query Match 37.0%; Score 2489; DB 9; Length 1280;
Best Local Similarity 42.1%; Pred. No. 5.9e-17;
Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

QY 94 EIKVNIISFGIMARVATKADILIMVISTCATAASTFORIMLY----- 137
DB 28 KKKKPTVSVFSMFRYSNMWDKLYNVGTLAIIHGAGLPLMLVGEWTDIFANAGNED 87
QY 138 -----QISYDEFF-----DELTKNLYEYVIGIGEVTVYVSTVGFIYGEHATOKI 184
DB 88 LMSNTNNSDINDITGFNNLEEDMTIRAYIYSGIAGVVAIYOVSWCLAAAGQIHKI 147
QY 185 REYVLESILRONIGYFDKLAGEVYTRITADTNLIQDISEKVGILTALATFYTAFTA 244
DB 148 RKQFPHAIMROBIGFVDHVDGELMTRLDDVSKINEGIDKIGMFMATFFGFIY 207
QY 245 YKWKIALICSTTVALTVMGGSOPIIKYKSKLDSYGAGTVAEEVISITRNATF 304
DB 208 FTRGKMLTLVLIATISPVGLSAVAKILISFTDELLAYAKAGAAVEVLAITFTVAF 267
QY 305 GTQDILAKQYEVHIDEAEKMGTKNOIYVGFIMIGAMFGLMNSYGLGFWMGSRFLVDGAVD 364
DB 268 GGQKELERYNKNLEAKRIGIKAITANISIGAAFLIITASYALAFWGTTLVLSGERS 327

Qy	365	VGDILVTLMLIGSFELGVSNNAOFTAAVAAAKITGGTIDROSPDLVSECKTIDH	4.2
Db	328	IGOVLTFFSVLLIGAFVSGAFSPSIEFANARAKAETIKRIDNKSIDSYSKSGHKPDN	387
Qy	425	FECHIELRNVKHIYPSRPEVTMEDVSLSPACKTALVPGSGSGKSTVVGIVERPEYMB	484
Db	388	IKGNLEFRNVHFSYPSRKEKILKGNLKVQSGOTVALVGNSSCGKSTVQLMRLYDPT	447
Qy	485	RGYVLDDGHDKDNLKMLAQOISLYSOEYVLEGTITIKINRGLIGTYENESDEKARE	544
Db	448	EGMVSVDGODIRITINVFLEIIEIGVSOEPLVETATIAENIRYG-----RENTYME---	499
Qy	545	LIEVAAKMAANADPTLPCGYEFTVNGOAREFLLSGGOKRIATARAAYADPKTILLDEAT	604
Db	500	-IEKAVVEANAYDPTIMLPIKFPDTLVGERKAQOISGGOKRIATARLYNPKTILLDEAT	558
Qy	605	SALDTKSEGVQAALERAABGRTTIVIAHRLSTIKTAHNIIVLVNCKIAEOGTHDELYDR	664
Db	559	SALDTSEAVAYQVALDKARKGRTTIVIAHRLSTVRNADYIAGDDGVIEKGNHDELMKE	618
Qy	665	GGAARKLIVENORINEKEDALDDADAEDLTNDIAKITASASDLDGKPTTIDRTGT	724
Db	619	KGILFKLVTVQTAGNEVE---LENA-----ADEKSELDALEMSNDNSRSLIRKRT	668
Qy	725	HKSYSAILSKRP-----PETTPKYSLTLLKLFASFNRPETIPMLIGVSVIAGGG	777
Db	669	RRSYRSGOADDRKISTEALDESIPVSPFRIMK-----LNTWTFPVVGVCAITINGEL	724
Qy	778	QPTQAVLYAKAISTLSLPESOYSKLRRHADFWSLMEFVVGIIIOFTQSTNGAFAVCSBR	837
Db	725	QPARAIIFFSKLIGVETIIDDPETK-RONSNMLFELFLAIGISFTIFFPQGFPGKAGEI	783
Qy	838	LIRABRSTARPTILRODIAEFKEENSGTMSFLSTERKHLSGVSGVTLGTIMSTLL	897
Db	784	LTKRLRTVWRSMKROVSNFDDPKNTGALTTRILNDAAYKGAIGSRILAYTTONIANL	843
Qy	898	GAALIIALAIKWLATVACISVNPVLLACGFYRFYMLAOFSORSKLAIEGSANFACEATSS	957
Db	844	GTGIIISFITYGMQTLTLLLAIVPLIIAAGVVENKMLSGQALXKDKRELEGAKATPAIEN	903
Qy	958	IRIVASITIREADWEIYHAOLDAGRTSLISVLRSSLIYASSOALVFCCVALGFWYGGTL	1011
Db	904	FRTYVSLTORCKEHEMVAQSLQVPRYRSILKAHIFIGTFESFQIOMMYFVACGFRGAAL	963
Qy	1018	LGHEHYDIPRFVCFSEILGASAGVSEFAADMDKAKRMAAEFRRLDRKPOQIDMSE	1077
Db	964	VAHKLMSFEDVLLFSAVVGAAAVGVSSFAVDYAKAKISAHNIIMITIEKRLDIDYST	1021
Qy	1078	EGEKLTEVEGEIEFRNVHFRYPTRPBQVPLRGDLTVKPGQYVALVPGSGCGKSTIALL	1133
Db	1024	EGLMPNTLEGNVGGEVFNVPTRPDIPVLQGLSTLEKKGQTLATLVGSSCGKSTVYOLL	1088
Qy	1138	ERFPDIAIGSLVDGKDISKLNINSTRSPISLYSOEPTLYOGTICKNIIILGIVEDDPRE	1199
Db	1084	ERFPDPLIAGKVLDDKCIKRLNQMURAHGIVISOEPIIDFCSIAENIAYGNSRVSQOE	1144
Qy	1198	FLIACADANIDYIEINSLPEGFMTVVGSKGMLSGGOKRQVATARALDLDPKILLDEAT	125
Db	1144	EIVAAKEANIHAFIESLPPKISTKYGDCKTQISGGOKRIATARLYNQPHILLDEAT	1201
Qy	1258	SALDSESEKVVQAAALDAARGRTTIVANRLSTIQRADYIYVFDQK IYESGHSLEYOK	1311
Db	1204	SALDTESEKVVQEALDAARGRTCIYIAHRLSTIQNADLIYVQNGRVENHGTHQOLLAO	1266
Qy	1318	KGRYVELVNIQS 1329	
Db	1264	KGIFYSMVVOA 1275	

RESULT 3
US-09-866-866A-2
; Sequence 2, Application US/09866866A

[illegible]

QY 778 QPQAVLYAKAISTLSPESQYSKLRHADFWMLFFVGIOTPTOSTGAAFPVCSER 837
 Db 725 QPFAAIFSKITIGVFRIDDEPK-RQNSMLFSLFLALGIIISITFLOGFFGKAGEI 783
 QY 838 LURASTAFRTLLRODIAFDKENSNGALSTFSLSTFKHLSGVSQVLTGLMTSTL 897
 Db 784 LTRRLKFRMFRSLRODVSPFDPKNTGALTTRLANDAAQVGAIGSRLAVTONTANL 843
 QY 898 GAAITIALIGMKLALVCISVPEVLACGFYRMYLAQFOSRSKLAVESGANSFACETS 957
 Db 844 GNGIISITFQMOQLTLILAIVPIIAIAGVEMKMLSGALKDKELEGAKIATEAIEN 903
 QY 958 IRRVASTLRDVMWEIYHQAQDQRTSLISVRSLSLASSQALVFCALGFWGGTL 1017
 Db 904 FRTVASTQOEKEHMYAQSLOVYRNSLRKAHIFGITSFTQAMMYFSYAGCFRGAYL 963
 QY 1018 LGHEHDIFREFVCFSEILFQAOSAGVFSFADPMCKAKNAAEFRRLPRKQDQNMWSE 1077
 Db 964 VAHKLMSFEDVLVFSAVYFGMAVGOVSFADYAKAKISAAHITMIEKTLIDYST 1023
 QY 1078 EGEKLEVEGELEFRVHRYTPRPOVYLKGLDLYVPGQVYALVPSGCGSTTIAL 1137
 Db 1024 EGIAPMTLBNVTFGEVFNYPTRDIPVLOGLSLEVKKGQTLALVSSGCGSTVQVL 1083
 QY 1138 ERYDVAISILVDGDISKLNINSTRSLVSOEPTLYOGTITENILIGVEDVPEE 1197
 Db 1084 ERYDVLAKRVLIDGKEIRLNVQMLRAHGIYSOEPIFLDCSIANIAVGDMSRYVSOE 1143
 QY 1198 FLTKACKDANIYDFMSLPEGEFTVYVSGKMGSGGOKORVATARALLRDKPILLIDEAT 1257
 Db 1144 EIVRAKKNIAHFISSLKPKISTKVGDKGTQSLSGOKORIALARLVKPHILLIDEAT 1203
 QY 1258 SALDSEKRYVQAALDAAARGTTIAVHRLSTIQRADVIYVEDQKIVESGTHSELVOR 1317
 Db 1204 SALDSEKRYVQAALDAAARGTTIAVHRLSTIQRADVIYVEDQKIVESGTHSELVOR 1263
 QY 1318 KGRYVELVMS 1329
 Db 1264 KGIYFSMVSQA 1275

RESULT 4
 US-09-769-097-4
 ; Sequence 4, Application US/09769097
 ; Patent No. US2002005128A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Kimberly Anne Brun
 ; APPLICANT: Richard James Chenery
 ; APPLICANT: Harna Ellens
 ; APPLICANT: John Anthony Feld
 ; APPLICANT: Lin Yue
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
 ; TITLE OF INVENTION: ENCODING RAT MDRIA AND SCREENING METHODS THEREOF
 ; FILE REFERENCE: GP-50009-C2
 ; CURRENT APPLICATION NUMBER: US/09/769,097
 ; PRIOR FILING DATE: 2001-01-25
 ; PRIOR APPLICATION NUMBER: 09/208,809
 ; PRIOR FILING DATE: 1998-12-09
 ; PRIOR APPLICATION NUMBER: 09/156,800
 ; PRIOR FILING DATE: 1999-09-17
 ; PRIOR APPLICATION NUMBER: US99/20770
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 1272
 ; TYPE: PRT
 ; ORGANISM: RATTUS RATTUS
 US-09-769-097-4

Query Match 37.0%; Score 2487.5; DB 10; Length 1272;
 Best Local Similarity 42.0%; Pred. No. 7.6e-177;

Matches 534; Conservative 226; Mismatches 461; Indels 49; Gaps 10;
 QY 90 KTLQLEIKVINSEFGMLRYATKMDILIMYSTICATAASTFORIMU----- 136
 Db 23 KKEKKKPAVSVLTFMRYAGWLDREFMLGLTALIIHIGLALMLVFECDMDSPANNG 82
 QY 137 -----YQIS--YDEYFDELTKNVLKYFVYLGIGEFYTVVSVGFIYGEHATOKIREY 187
 Db 83 NNRSMSFYNTDIAKLEDEMTYAYYTGIGAGVLIIVAYIQVSTWCLAAGROIHKIROK 142
 QY 188 YLSILRONIGYDCKGAGVETTRITADTNLIDQISSEKGLTLTATFVTAAYK 247
 Db 143 FFAIMNOEIGMPDVHDVGEINLRLDQVSKINEGIGDKTGMFQMAATFFGGFIIGFR 202
 QY 248 YMKLALICSTTIALVLTMGGSQFIKYSKSLDSYGAGTVAEEYISSIRNATFAGTQ 307
 Db 203 GMLTLVLIATISPLVGLSAGIMAKLISSEFDKLOQVAKAGNAVEYLAIRIVAFGG 262
 QY 308 DKLAKOYVHLDAEKGKTKNOIVMGFMIGAMFGLMSNYGIGFMGSRFLVDGAVDGD 367
 Db 263 KKELEERYNNMLEBAKRIGIKAITANISMGAAFLLIYASVALAFMYGTSLVISKEYTIG 322
 QY 368 ILFVLMALIGSFLGNSPNAQAFINAVAAAKITGITDROSPIDRPSMEGTLDHFE 427
 Db 323 VLTVEFSLVIGAFSVQASPNIEAFANARCAAEVFSIIDNKSIDSFSKGRKPNIOG 382
 QY 428 HIELRWVKHIYPRPEYTVMEDVLSMPAGKTALVGPSSGSGSTVAGYERFEPVGR 487
 Db 383 NLEFKNIHFSYPRKQVOILKGLNLKLVKSGQVYALVNSGCGSTVQVLQRLYDTEGE 442
 QY 488 VLLDHDIDKDLNRLTROIISLVSOEPLVFGTTIYNIRHGLIGTYENESDKVRELIE 547
 Db 443 VSIDGDIRTINRYLREIIGVYSOEPLVFATIAENIRYG-----RENTYME----- 493
 QY 548 NAKMANADEFTALPEGETVNGORGLISGOKORIVATARYSDPKTILLIDEATSL 607
 Db 494 KAYKEMNAIDFLMKDHPKEDTLVGERGAQSLSGOKORIALARLVKPHILLIDEATSL 553
 QY 608 DTKSEGVQAALERAAGRTTIVIAHRLSTIKTAHNIYVLYNGKIAEOGTHDELVDGGA 667
 Db 554 DTSEAVVQAALDAAKAREGRTTIVIAHRLSTVRADVIAGPDGVYIEQGNHDELMKEGI 613
 QY 668 YKRLVDAQORINEKEA--DALEPADADLTNADIAKTAASSASDLDKPTTIDRTGHN 725
 Db 614 YKRLVMTQAGNIEELGNACSEKSDID--NVDMS--SKDSQSSILIRRSRFSKIRGPH 668
 QY 726 KVSASAILSKRPETPKYSIMTLTKFVASFNPEPYMLIGVFSVGGGPOQAVLY 785
 Db 669 DQDGLSTKEALDDVPPASFWRLK-----LNSTEMPYVGVFCALINGGLDPAFSTIF 724
 QY 786 AKAISTLS--LPESQSKLRHADFWMLFFVGIOTPTOSTGAAFPVCSERLIRRA 842
 Db 725 SKYVGFTKNDPFEIQ-----RQNSMLFSLFLALGIIISITFLOGFFGKAGEILTREL 780
 QY 843 RSTAFRTLLRODIAFDKENSNGALSTFSLSTFKHLSGVSQVLTGLMTSTTGAIT 902
 Db 781 RYMWFSKMLRODISWPDDEKNTGALTTRLANDAAQVGAIGSRLAVTONTANLSTGII 840
 QY 903 IALAIGMKLALVCISVPEVLACGFYRMYLAQFOSRSKLAVESGANSFACETSIRTYA 962
 Db 841 ISLITGMQLTLILAIVPIIAIAGVEMKMLSGALKDKELEGAKIATEAIENRTVY 900
 QY 963 SLTREDVWEIYHQAQDQRTSLISVRSLSLASSQALVFCALGFWGGTLLGHNE 1022
 Db 901 SLTREQKFETMYAQSLOIYRNLAKKAHFGITFSTQAMMYFSYACRFDAYIYARBL 960
 QY 1023 YDIFRFVCFSEILFQAOSAGVFSFADPMCKAKNAAEFRRLPRKQDQNMWSEGERL 1082
 Db 961 MFEENVLIVFSAIVFGAMVGOVSFADYAKAKISAAHITMIEKTLIDYSTGLKP 1020
 QY 1083 ERYEGERFRVHRYTPRPOVYLKGLDLYVPGQVYALVPSGCGSTTIALERFD 1142
 Db 1021 NMIEGNVKNFGVWFNYPTRDIPVLOGLSLEVKKGQTLALVSSGCGSTVQVLLERFYD 1080


```

; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent version 3.0
; SEQ ID NO 4
; LENGTH: 1280
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-866-866A-4

Query Match      36.9%; Score 2480; DB 10; Length 1280;
Best Local Similarity 42.0%; Pred. No. 2,8e-176;
Matches 534; Conservative 229; Mismatches 449; Indels 60; Gaps 9;

QY 94 ERIKYNISFGIMRATKMDILIMVISTICALAASTFORIMLY----- 137
DB 28 KKKKPTVSFVSFMRYSNMJLKLMMVGTLLAIIHGALPLMLVFGEMTDIFANAGNLED 87
QY 138 -----QISYDEFEY-----DELTKNVLVYFYLIGIGEFVTVYSTVGFITYGEHATOKI 184
DB 88 LMSNTNRSDINDTGFPMLEDMTRVATYISGIGAVVAAYIYQVSEFCLAGROIKHI 147
QY 185 REYVESLIRONIGYFDKIGAGEVTRITADTNLIQDISERVGILTLTALATFVAFI 244
DB 148 RKQFHAIRWROEIGWFDVDELNTRLTDPVSKIWEIGDKIMGFQGMATFEFGIY 207
QY 245 YKWKMLILICSTVALVLTMGSGGSQIFIKYSKSLDSYGAGGYAEVVISIRATAF 304
DB 208 FTRGMKLTLLVLAISPVLGLSAVMAWKILISFTDELLAYARAGAAVEVLAIRVIAF 267
QY 305 GTODKLAEYEHVDAEERKGTNOIYMGFMIGAMEGLMYNKGFGMSREPLVDAVD 364
DB 268 GGQKKELEBRYNKNLEAKRIGIKKAITANISIGAEFLIYASTALAFYGTLLVLSGETS 327
QY 365 VGDILTVMALLIGSFSGVNSPMAQFTNAVAAAIFGTIDROSPLDPYSNEGTLDH 424
DB 328 IGOVLTFFSVLIGAFVSQAQSPSIEAFANRGAAYEIRKIDNKRPSIDYSKSGHKPN 387
QY 425 FEHGELENNVKHIYPSRPVTVMEQVSLSMRPAKTTALVPGSGSKSYVGLVERMYM 484
DB 368 IKNLEFRNVEHSYSPSRKVKILKINLKVQSGQVALVNGSGCKSTVOQMORLYDPT 447
QY 485 KGVLLDHDIKDLNRLRQOISLVSQEPVFGTTIYKINRHLIGTVESEDEKYE 544
DB 448 EGMVSVDDGODIRTIIVRFLREITIGVSOQEPVLFATTIAENIRYQ-----RENVYDDE--- 499
QY 545 LIENAKANAHDFITALPEGETVNGORGFLLSGQKORALAIARAVVSDPKILLDEAT 604
DB 500 -IEKAVKEANAYDFIMKIPHKFDLVGERGQSLSGQKORIALAIARALVNPRIILLDEAT 558
QY 605 SALDRKSGVVOALERAAGRTTIVIAHRTSTIKTAHVIVLVNGKIAEGSTHEDVDR 664
DB 559 SALDRESEAVVOALDKRKGRTTIVIAHRTSTIKTAHVIVLVNGKIAEGSTHEDVDR 618
QY 665 GAVYKRLVEAORINEOKRADALEDADAEDLTNAADIAKIKITASASSDDLGKPTTIDRTGT 724
DB 619 KGIYFKLVTMGTAGNEVE---LENA-----ADESKSEIDALEMSNDRSLSLRKST 668
QY 725 HKSVSSAILSKRP-----PETTPKYSLWTLTKFVASFNREPIPYMLIGLVFSLVAGGG 777
DB 669 RRSYVGSQAOBRKLTSTKEALDESIPVSFWRIMK---LNLTEPMYFVGVGVCALINGL 724
QY 778 QPTQAVLYAKAISTLSLEPSQSKIRHADQWLSLMEFVVGIIQFTOSTNGAAPVCSER 837
DB 725 QPAVALISKTIQVTRIDDEPK-RONSNNFSLFLALGIIISITTFLOGTFKAGBI 783
QY 838 LIRARASTAFITLRODIAFQKENSATGALTSFLSTENKHLSGVGYTLGILMTSTTL 897
DB 784 LTKRLRYVFSMLRKQDVSWEDDPKNTITGALTTRILANDAAOVKGAIGSLAVITONIANL 843

```

```

QY 898 GAIIITAIAGMKIALVCISVPEVLLACGFYRPMYLAOFOSRKLAYEGSAPACEATSS 957
DB 844 GTGIIISIFYGQMLTLLLAIVPIIALAGVEKMKLSGALKDKKLEBAGKATATAIEN 903
QY 958 IRTVASLFRBDVWEIYHAOLDQGRFTSLISVRLSSLIYASSQALVFCVALCFWYGTLL 1017
DB 904 FRTVASLTOEOKFEHMYAOSLQVPRNLSLKAHIFGTFPSFOQAMFYFACGFRGAVL 963
QY 1018 IGHHEYDIFRFVCFSEILIFGASQATVSEFAPDKAKANAAAEFRRLDRKPOIDNMSE 1077
DB 964 VAKKLMSFEDVLVFSAAVVGAAVQVSSFADVAKAKATISAAHITIMITEKPLDLSYST 1023
QY 1078 EGEKLETEGELEFRNVRHPTPEOPVLRGLDLTVKGGQYALVPGSGCSTTIAL 1137
DB 1024 EGMPTLELGNVTFGVSVEYNTPRPIPLQGLSLEVKKGQTLALVSSGCKSTYVOLL 1083
QY 1138 ERPYDAISLIVDGDISKLINISYRSLSVLSQEPPLLYOGTIKENILGIVEDDPEE 1197
DB 1084 EREYDPLACKVLLDCKEIRLNVQWMLRAHGIYVQEPILFDCSIABNIAYGDNSRVSQE 1143
QY 1198 FLTKACKDANIYDEIMSLPEGFNTVYSGKMGMSGQKORVALIARLLDPRKILLDEAT 1257
DB 1144 EYRAKKEANIHAFTESLNPKNSTKVGDKGTQLSGQKORIALARLVQPHILLDEAT 1203
QY 1258 SALDSEKRVVOALDPAARGRTTIVAHRLSTIQKADVIYFQDGIYESGTHSELVOK 1317
DB 1204 SALDSEKRVQDALDKRAGRCIYIARLSIQNADLIYVQNRVREHGHQOLLQ 1263
QY 1318 KGRYFELVLOS 1329
DB 1264 KGIYFSMVSVOA 1275

```

RESULT 7

```

US-09-769-097-2
; Sequence 2, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Eilens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1272
; TYPE: prt
; ORGANISM: RATTUS RATTUS
US-09-769-097-2

```

```

Query Match      36.8%; Score 2475.5; DB 10; Length 1272;
Best Local Similarity 42.0%; Pred. No. 5.9e-176;
Matches 533; Conservative 225; Mismatches 463; Indels 49; Gaps 10;

QY 90 KTOLEIKYNISFGIMRATKMDILIMVISTICALAASTFORIMLY----- 136
DB 23 KKKKPKPAYSVLNMFRTAGLDRFYMILGTLAIIHGIALPLMLVFGMDTSFANVG 82
QY 137 -----YQIS--YDEYDELTKNVLVYFVYVIGIGEFVTVYSTVGFITYGEHATOKIREY 187

```

Db 83 NNRSMFYNAADTYAKLEDEBMAITYYYTNGCGVLIVAIQVSLMCLAGROIHKIRK 142
 QY 188 YLESIIRONIGYEDKLGAGEVTRITADTNLIDGISEKGLTTLTALAFVFAIYAVR 247
 Db 143 FFAIINNOIGEMDVADVELNRLDDVSKINEIGDKIGMFQOMATPFEGFIIGTR 202
 QY 248 YMKALICSTYVALVLMGGGSOFTIKYSKSLDSTYGAGGVAEEVYSSIRNATAFGTQ 307
 Db 203 GMLTFLVILAIISFVLGSLAGIMAKILSSFTDKELQAVAKAGAAVEVLAIRFVIAFGGQ 262
 QY 308 DKLAKGYEHLDAEKMGTRKNQIVMGFMIGAMGLMYSNNGLGFMWGSRLVYGAVDVG 367
 Db 263 KKELEFRYNNNLEAKRGLGKALTANISMGAAFLLIYASTALAFWGTSLVISEKTYITIGQ 322
 QY 368 ILTVLAILIGSFSLGNVSPNAQFTNAAVAAKIFGTIDROSLDPDYSNEKTLDFHEG 427
 Db 323 VLTVFVSVLGAFVSQASNIIEAFANAGAAEVEYSIIDNKSIDSFSKGRPNIOG 382
 QY 428 HIELRNKHLIYSPRPVYVMEVSLSPAGKTTALVPGSGSKSTVGLVREYFMPVGT 487
 Db 383 NLEFKNIHFSYPRKDVOJILKGLNLKYKSGQVALVNGSCGSKSTVQLQRLYDPIEGE 442
 QY 488 VLLDGDIDNLNRLRQOISLVSOEVLFGTTIYKNIHRLGLGTRENSDEKVRLE 547
 Db 443 VSIDGODIRITNRYLREIIGVSOEVLFAITIAENIRG-----RENTYME---IE 493
 QY 548 NAKMANADFTALPEGETYNGVGRGFLSGGOKORAIARAIVSDPKILLDEATSA 607
 Db 494 KAVKEMAYDFIMKLRHKFDVLGKGAQLSGGOKORAIARALVNRPKILLDEATSA 553
 QY 608 DTKSEGVQALERRAEGRTTIVIAHRLSTIKTAHNIIVLVNCKIAEQTHDELVDNGA 667
 Db 554 DRESEAVVOAALDKAEGRTTIVIAHRLSTVRNADYAGDGVYIEQGNHDELMREKGI 613
 QY 668 YRVLVQAOIRNEQKA--DALEDAEDLTMADIKIKTSASSDLDGKPTIIDRGTH 725
 Db 614 YFELVMTQAGNIEELGNECESKSDGID--NVMS--SKDSQSSLLRRSTKIRGPH 668
 QY 726 KVSASAILSKRPETPKYSIMTLKFKVASFNRPETPYMLIGVFSVLAGGQPTQAVL 785
 Db 669 DODGELSTKEALDDVPPASFWRLK----LNSTEMPYFVVGFCALINGLOPABSITF 724
 QY 786 AKAISTLS---LPESQSKLRHDADFWSLMFVYVGIIOFTVSTNGAFAVCSERLIRA 842
 Db 725 SKVAVGTFKNDPEIQ-----RONSNFLSLFLILIGISFTFLOGTFPKAGBIILKRL 780
 QY 843 RSTAEFTIIRODIAFEDKENSFGALTSPLSTFKHLSVSGVTGLGIIIMTSTTGAAIT 902
 Db 781 RYWFKSMRLKODISWDDPKNTTGALTTRLANDAAOVKKGWGSRLAVITONIANLGTII 840
 QY 903 IALAIKMLALVCISVVPVLLACGFYFYMLAOFOSRSKLAEGSANFACETASIRTV 962
 Db 841 ISLIYMOQLTLLLAIVPIAINGVEMKRLSGOALKDKKELEGSGKIAEALENFTV 900
 QY 963 SLTREBOWEYIAQLDAOGRTSLISVNLSSLLYASSQALVFCVALGFYGGTLLGHE 1022
 Db 901 SLTREKFEEMYQOSLOIPRNALKRAHVGIFTFSEFOAMMYFYAACFRDVALVAREL 960
 QY 1023 YDIFRFVCESELLFGAQSAGTFESFAPDKMKANAAERLDFDRKPOIDNNSSEBEK 1082
 Db 961 MPEENLVFSAIVFAMAAGOVSSFPADAKKVASHSIRIITEKIPEDISSTEBELKP 1020
 QY 1083 EYVEGEIEFFNVHFRPTREPOVLRGLDITVPRGQVALVPGSGSKSTTIALLEFRYD 1142
 Db 1021 NMLEGVAKFNGVMEFNPTRNIPVLOGLSLEVKKGOTIALVSSGCKSKSTVQLLEFRYD 1080
 QY 1143 AIMGSLVLDGKDISKLNINSYRSFSLVSOEPLLVOGTIKENILGLIVEDVPEELIKA 1202
 Db 1081 PMAGTFLDGEKEIKOLNVOLRAHLGIVSOEPLITFCSISIENTIAYGDNSSHYVSKEIYKA 1140
 QY 1203 CRKANITYDEFTMSLPEGFNTVYVSGKMLSGGOKORAVIARALLRDPKILLDEATSA 1262
 Db 1141 AKBANIHOFIDSLPERKYNTRVSGKGQOLSGGOKORAIARALVNRPKILLDEATSA 1200

QY 1263 ESEKVOAALDAAARGTTIAVAHRLSTIQKADVIYVDOGKIVESGTHSELVOKKGRY 1322
 Db 1201 ESEKVOAALDKAREGTCVIAHRLSTIQNADLIIVIQNGVKEHSTHQOLLAQKITYE 1260
 QY 1323 ELVNLQSLGK 1332
 Db 1261 SMVSVQAGAK 1270
 RESULT 8
 US-09-866-866A-8
 ; Sequence 8, Application US/09866866A
 ; Patent No. US20020102244A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sorrentino, Brian
 ; TITLE OF INVENTION: A method of identifying and/or isolating Stem Cells
 ; FILE REFERENCE: 1340-1-021CIP2
 ; CURRENT APPLICATION NUMBER: US/09/866,866A
 ; CURRENT FILING DATE: 2001-08-30
 ; PRIOR APPLICATION NUMBER: 09/584,586
 ; PRIOR FILING DATE: 2000-05-31
 ; PRIOR APPLICATION NUMBER: PCT/US99/11825
 ; PRIOR FILING DATE: 1999-05-27
 ; PRIOR APPLICATION NUMBER: 60/086,988
 ; PRIOR FILING DATE: 1998-05-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 8
 ; LENGTH: 1276
 ; TYPE: PRF
 ; ORGANISM: Mus musculus
 US-09-866-866A-8
 Query Match 36.8%; Score 2475.5; DB 10; Length 1276;
 Best Local Similarity 41.4%; Pred. No. 6e-176;
 Matches 527; Conservative 230; Mismatches 466; Indels 49; Gaps 6;
 QY 90 KTOLEELKVINISFGMLRATKMDILIMVSTICAIAMASIFQIMX----- 137
 Db 23 KKEKERRKPAVSVLTMRVYAGWMDRLVMTGLAAIIGHVALPMLMIFGDMTDSFASVG 82
 QY 138 -----QISYDFEYDELTKNVLVYVVLGICEFYTVSVYGTGTGEHATOK 183
 Db 83 NVKSNSTNMEADKRAFALEEBMTTYAYTYIGAGVLIYAIQVSPWCLAGROIHK 142
 QY 184 IREYIESTLRONIGYEDKLGAGEVTRITADTNLIDGISEKGLTTLTALAFVFAI 243
 Db 143 IROKFFHAIINNOIGEMDVADVELNRLDDVSKINEIGDKIGMFQOMATPFEGFI 202
 QY 244 AYVYKMKALICSTYVALVLMGGGSOFTIKYSKSLDSTYGAGGVAEEVYSSIRNATA 303
 Db 203 GFTRGKRLTVILAIISFVLGSLAGIMAKILSSFTDKELHAAKAGAAVEVLAIRFVIA 262
 QY 304 FGTQDKAKOYEVLDAEKMGTRKNQIVMGFMIGAMGLMYSNNGLGFMWGSRLVYGAV 363
 Db 263 FCGOKKLEERYNNNLEAKRGLGKALTANISMGAAFLLIYASTALAFWGTSLVISEK 322
 QY 424 HFEHIELRNKHLIYSPRPVYVMEVSLSPAGKTTALVPGSGSKSTVGLVREYFMP 483
 Db 383 NIOGNLEFKNIHFSYPRKEVQJILKGLNLKYKSGQVALVNGSCGSKSTVQLQRLYD 442
 QY 484 VRGTVLLDGDIDNLNRLRQOISLVSOEVLFGTTIYKNIHRLGLGTRENSDEKVR 543
 Db 443 LDGKVSIDGODIRITNRYLREIIGVSOEVLFAITIAENIRG-----RENTY 493
 QY 544 ELIENAKMANADFTALPEGETYNGVGRGFLSGGOKORAIARAIVSDPKILLDEA 603

```

Db 494 DEIEKAVKAAVDFIMKLPHEFDLIVGERGAHVGSGGOKORAIANALVBNPKIILLDDA 553
Qy 604 TSALDTKSEGVVQAAALERAEGRTTIVIAHRLSTIKTANIVYLVNGKIAEGTHDELYD 663
Db 554 TSALDTSEBAVVOALDKAREGRTTIVIAHRLSTVBNADVINGFDGCVIVEGDNHDELR 613
Qy 664 RCGATRKIVLEAORINEOKRADALEDADADLNADIATKKTASSASSDIDGKPTTIDRIG 723
Db 614 EKGITFKLMTQTAGNEILGNBCKSKDEIDNLDMS---SKDSGSSLLRRRSTRKSIG 670
Qy 724 TRKSVSAILSRKPEPTKYSLMTLLKFVASFNRPEIRPEIIGLVSVIAGVGOPTOAY 783
Db 671 PHDQBRKSTKALDEVDVPASFWRILK---LNTSEMPYFVVGICAILINGIDAFESY 726
Qy 784 IYAKAISTLS---LPESQYSLKRDADFWMSLMEFVVGIIOFTQSTNGAFAVCSERLIR 840
Db 727 IFSKVGVFTNGGPETQ---RONSNLFSLLFLILGISFTFELQGFPGKAGEILRK 782
Qy 841 RARSTAFRTILRODIAFPEKENSNGALTSPLSTERKHLISGVSVTLGITLMTSTTLGA 900
Db 783 RLRVAFKSMLRQDVSWEDDPKNTGALTTLANDAAQKGAISRLAVIFONIANILGTG 842
Qy 901 IIALAIGKIALVLCISVVPVLACGFYFVYMLAQOFOSRSLAYEGSANFACEATSIRT 960
Db 843 IILSLYGMQLLILLALVPIIAGVEMKMLSGALKDKELBESGKIATEPAIDENFRT 902
Qy 961 VASLFRERDWEIYHAQDLDAQGRSLISVLRSSLLYASSQALVFECVALGFVYGGTLLGH 1020
Db 903 VVSLRQGFETFTYMAOSLIQIPYRNAMKKAHVFGITFSSFOAMMYFYACFRGAVLYQ 962
Qy 1021 HEYDIFRPFVVCSEILFGQASAGYVSPAPDMGKAKNAAEERRRLFDKRPQIDNMSEEB 1080
Db 963 QLMTEENVLVSAVFGMAAVGVSSFPADAKATVSAHSIRIETKEPDISYSTOGL 1022
Qy 1081 KLEVEGEIEFRNHFRTPEQPYLRGLDLTVKRGQVALVGPSCGCKSTTILLERF 1140
Db 1023 KPNMLEGNVQFSGVFNFTFRSIPYLOGSLLEVKGQTLAVLGSGCKSTVYQVLEFR 1082
Qy 1141 YDAINGSLVLDKDISKLNINSRFSLSVSOEPTLYOGTIKENILGLIVEDVPEEFLI 1200
Db 1083 YDPMAGSVFLDKETIKQLQVLMRAOLGIVSOEPLIFDCSIENATVAGDNSRVSYEEI 1142
Qy 1201 KACKRANIVDFIMSLPEGRNVTYVSGKGLSGGOKORVAAIALRDPKILLDETSAL 1260
Db 1143 RAAKAMNHQFIDSLPDKNTFRYGDQGLSGGOKORALAAIALVAKQPHILLDETSAL 1202
Qy 1261 DSESEKVVQAAALAAARGTITVIAHRLSTIQADIVYVFDQGIYESGTHSELYOKGR 1320
Db 1203 DTESKVVQEAALDKAREGRTTIVIAHRLSTIQADIVYVYONGKYEKRGTHOGLAOKCI 1262
Qy 1321 YVELVNLQSLGK 1332
Db 1263 YFSMVSVOAGAK 1274

```

RESULT 9
US-09-866-866A-6

```

; Sequence 6, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866, 866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584, 586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086, 988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27

```

SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 1276
TYPE: PR
ORGANISM: Mus musculus
US-09-866-866A-6

Query Match 36.5%; Score 2453.5; DB 10; Length 1276;
Best Local Similarity 40.9%; Pred. No. 2,6e-174;
Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;

```

Qy 90 KYOLEIKVNIISFGIMRATKMDIILVYISTICAIAASTFORIML----- 136
Db 23 KKEKKKKPAGVGFGRADWLDKLCMTLIGTLAIIHGTLLPLMLVGNMTDSPTKAE 82
Qy 137 -----YQSYDEYDELTKNVLFEYFLIGEFVYVYVSGVITYGERA 180
Db 83 ASLISITNOSGPNSTLIISNSLSLEEMAIYAYTIGICAGVLIYVIOVSLMCLAGKQ 142
Qy 181 TOKIREYLESILRONIGYFDKLGAGEVTRITADTNLIDGISEKVLTLTALATFVA 240
Db 143 IHRKQKFHALINNGEIGFVDHVDGELNTRLDVSKINDIGDKIGMFQSITFFLAG 202
Qy 241 FIITAYKYKIALICSTVALVLTVMGGSQFIKISKSLDSYGAGTVAEVISSIN 300
Db 203 FLIGFISGKLLVLIYVPLIGLSALMAKVLTSPTNKELAYAKAGAAVEVIAAIRT 262
Qy 301 ATAFGTQDLAKOYEYHDEAEKQKNOIYMGFMIGAMEGLMYSNYGIFMMGSRFLVD 360
Db 263 VIAFGQOKELERYKNLBEAKNVGIIKAITASISIGIYLVLYVSYALAFWYGISLVIS 322
Qy 361 GAVVDGILITVLMALIGSFSLSGNVSPNAQAFNAAAKIFGTIDROSPDPPSNEBK 420
Db 323 NEXSIDEVLTFEFSILGLFSGIHLAPNIEAFANAGAFELFKIIDNPSIDSTKCY 382
Qy 421 TLDHPEGHLELNKVIYSPREYVYMEDVSLSPAGKFTTALVGPSCGCKSTVGLVERE 480
Db 383 KPDISLNGNLEFNVHFNYSRSEVOILKGLNLYKVSQGTVALVNGSGCKSTVQIMQRL 442
Qy 481 YMPVRGTVDLGDHIDKIDMLRMLROQISLVSOEPLYFGTITKNTFRHGLIGRYENESD 540
Db 443 YDPLEGVASIDODITFTVIRKILREITIGVSOEPLYFATITLIENTNYG-----RD 493
Qy 541 KYRELIENAKKAMAHDFLTALPEGETVNVGQFELLSGGOKORAIANAVYSDPKILL 600
Db 494 YTMDEIEKAVKAAVDFIMKLPHEFDLIVGERGAQLSGGOKORAIANALVBNPKIILL 553
Qy 601 DEATSLDTPKSGVVOALERAEGRTTIVIAHRLSTIKTANIVYLVNGKIAEGTHDE 660
Db 554 DEATSLDTPSEBAVVOALDKAREGRTTIVIAHRLSTVBNADVINGFDGCVIVEGDNHDE 613
Qy 661 LVDRGAYRKIVLEAQ-RINEOKRAD---ALEDADADLTNADIATKKTASSASSDIDOK 715
Db 614 LMRKGIYFKLVMQTQRGHEIEPGNNAVYOSQDITDASELTSESKSPLIRSIYSVHAK 673
Qy 716 PTTIDRTGTHKSVSAILSRKPEPTKYSLMTLLKFVASFNRPEIRPEIYMLIGLVSVIAG 775
Db 674 QOQERRLSKAEAYD-----EDVPLVSFWRL---NLMSLMPYLVGLVLAING 720
Qy 776 GGQPTQAVIYAKAISTLSPEQYSLKRDADFWMSLMEFVVGIIOFTQSTNGAFAVCS 835
Db 721 CIOPVFAIVFSRIYGVFSDDDHETK-RONCNLFSLFELVLMGLISVYVFFQGFPGFKAG 779
Qy 836 ERLIRARSTARFTILRODIAFPEKENSNGALTSPLSTERKHLISGVSVTLGITLMTST 895
Db 780 ELITRKVRIMVYKSMLRQDISFDDHKNSTGSLTTRILSDASSVGAMCARLAVTQNYA 839
Qy 896 TGAALIIALAGWKIALVLCISVVPVLACGFYFVYMLAQOFOSRSLAYEGSANFACEAT 955
Db 840 NLGTGVIISLVYGMQUTLILLVITPLIYVIGLIEKKLSGQALKKQKOLEISGKIATEAI 899
Qy 956 SSIRIVASLIRERDWEIYHAQDLDAQGRSLISVLRSSLLYASSQALVFECVALGFVYGG 1015

```



```

? SEQ ID NO 5
? LENGTH: 1222
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: NO. US20020037522A1e
? LOCATION: (230)..(230)
? OTHER INFORMATION: Xaa at position 230 represents any L amino acid
? OS-09-873-409-5

```

Query Match	33.6%	Score 2258.5	DB 10	Length 1222
Best Local Similarity	40.3%	Pred. No. 8	8e-160	
Matches	487	Conservative	225	Mismatches 434
				Indels 61
				Gaps 10

QY	153	LYEFLVLCIGEEVWYVSTVVFVYGEHATQIKREYLESTIRONICYEPFKLAGETVETI	212
Db	47	LYVYGIVAAVLREYIDISLMTITTAARQPIRIKQPFHSHVLMODICWPEOSDGLINTPM	106
QY	213	TADTNLIODGISEKVGLTALTAPVTAFTIAYVYKWLALICSSPTVALVLTMGGSQF	272
Db	107	T-DIDKISDGIKRIALLFONMSTFSIGLAVGLKWKMLTLYLTSPITMAAASRM	165
QY	273	IIKRSKLSYGGAGGVVAEVSISIRNATAPFOOKLAK-----	312
Db	166	VISLTSKELSNYSKAGNAVEEVLSSIRPYVAFAOKEIKORSPLINTIRYAFEPOMLL	225
QY	313	-----QYEVHIDEAEKWTGNQIYGMWIGAMGLMTSNYGLGFWMGSREPLYDG--AV	363
Db	226	SCVLXFPRYONLKDADDFGIKRTIASKVSIGAVYFPFMNGTGLAWYGTSLILNBPXY	285
QY	364	DVSGILVLMWAILIGSPISLGNVSPNQAFTNNAANAARKFQTRIDRSPIDRPSNGKTIJD	423
Db	286	TIGVTLVAFPSYHSSICIGAAVPHETETIAGAAPHFVYIDKPSIDNSTAGVXPE	345
QY	424	HFEGHIELRWKHLIYPSRPEVYMEVDSIMPAKTTALVGBSGSGSKTVYGLVERFYMP	483
Db	346	SIEGTVEFKNVSPFSPSFIKILKGLNLIKSGEFTVYALVGLNGSGSKTVYQDLQRLYDP	405
QY	484	VRGVVLDDGHDIKDNLNRMLROOISLVYSQBPVLFEGTTIKYKRIHGLIGTKYENESDYKR	543
Db	406	DDGFIWDENDIRLNVRAHRIHIGVYSDPVLFEGTTISNIRKYG-----RDVYTD	456
QY	544	ELIENAKMANADHETALPEGEYETWVGCGFLISGSGOKORIAARAVSDPKILLIDDA	603
Db	457	EEHMERARANAADFIEMEPNKNTLVYGEKGAWSGQORIALAVALRNKRILLIDDA	516
QY	604	TSALDPTSEGVQAALERRAAEGRTTIVIAHRSTITYAHNIYVLVNGKTAEOGTHDELVD	663
Db	517	TSIADDSKSAVQAALAEKASKGRTTIVVAHRSTISADLIYTLKGMIAEAGHAELWA	576
QY	664	RGAVRYLVEAORINEOKEDALDEDADDELTAADIAKIKTASSSASDLDGKPTTIDRTG	723
Db	577	KRGITYLSVWSODIKKADDEMSMTYTEKNTS--LPLHSYKSIKSDP-----IDK--	626
QY	724	THNSVSALLSKRPETTPRYSLMTLTKLVASFNRPELPIYMLIGLVFYSIAGGQPTQAV	783
Db	627	AEBSTQKREIS-----LPEVSLIKLIK-----LNKPEMPFVYLGTLASVLTNGTVHVESI	676
QY	784	LYNAKISTLSPEQSOKLBDADPFWSIMFVYVGIQIOTOSTNGAAFAVCEBRILRRAR	843
Db	677	IFAAIITYMFG--NNDKTYTLHDAEYISMLFVLLIGVLCFSPYTMQGLFYGRAEYITLMRIR	734
QY	844	STAFRTILRODIAFFDEKEENSTGALTSFLSTERNKHLISGVSGVTLGTLMTSPITLAAIIT	903
Db	735	HLAFKMLYODIAMFDEKREKSTGGLTILALDIAQIQAGATGSRIGVLTIONATINMGLSVIT	794
QY	904	ALAIQWKLALVCISVVPVLLACGFYFPYMLAOFQSSKSLAYEGSANFACCATSSIRTVAS	963
Db	795	SFYIYGMETPLILISAPVLAVLTGMITETAAMTGFAKNDOKELHACKIATEALENTRITYS	854
QY	964	LTTERDOWEYTHAQLOAOGRTSLISVLRSLSLYASSQALYFECVALLGFWYGGTTLIGHNEY	1022
Db	855	LTRKAPAEQWTEKLOTHNNISKKAQOIGCSTAFSHATYFRAVAGFFFGVALLQAGRM	914

QY	1024	DIFFEFCESILLGAQSACTIVESFADPMCKANNAAEFRRLIDRKPQIDMNSSECKLE	1083
Db	915	TPEGMFIVFTAIATVAGNAIIGKTLVLAPEYSKASGAHHFALLEKPPIDNSQSGKKRP	974
QY	1084	TVEGEIERUNVHERFYPTRPBQPVLRGLDILVPEKQYVALVCGSGCKSTTILLERFYDA	1143
Db	975	TCEGHLERREVSFFYPCRPDPVFTILRGJLSTIERGKTVAVGVSSGCKSTSYVLLQRLYDP	1034
QY	1144	IAGSLVDGKDISKLNINSTRSFLSVSOEPTLYOGTITENILLGIVEDDYEEFLIAC	1203
Db	1035	VQCGVLPDGVAKELANQMLRSQIAIATVQEPVLFNCISIAENIATVGDNSRVYPLDEIKBA	1094
QY	1204	KDANITVDFIMSLPGEFTTVVSGSGKMLSGGQKORVAIAALLRDKILLDEBATSALDSE	1263
Db	1095	NAANIHSEIEGLPEKTYITVOGLGAOLSGGQKORLAIARALLQKPKILLLEATISALDND	1154
QY	1264	SEKVVQALDAALAAAGPTTIVARHLSTICKADVIYFEOGKIVESGTHSELVOKGGRYVE	1323
Db	1155	SEKVVQALDAARTGCTLVATVHRLSHLQIADLVILHNGKIKEGTHQOELLRNPDIFYK	1214
QY	1324	LVNLOSL 1330	
Db	1215	LVNAGSV 1221	

```

1 RESULT 12
2 US-09-749-340-6
3 : Sequence 6, Application US/09749340
4 : Patent No. US20020037920A1
5 :
6 : GENERAL INFORMATION:
7 : APPLICANT: Prichard, Roger K.
8 : APPLICANT: Xu, Ming
9 : APPLICANT: Ribeiro, Ana Paula
10 : APPLICANT: Blackhall, William J.
11 : APPLICANT: Beech, Robin N.
12 : APPLICANT: Molento, Marcelo
13 : APPLICANT: Liu, Hao Yuan
14 : TITLE OF INVENTION: Methods for Detecting and Reversing
15 : TITLE OF INVENTION: Resistance to Macrocyclic Lactone Compounds
16 : NUMBER OF SEQUENCES: 12
17 : CORRESPONDENCE ADDRESS:
18 : ADDRESSEE: American Home Products Corporation
19 : STREET: One Campus Drive
20 : CITY: Parsippany
21 : STATE: New Jersey
22 : COUNTRY: United States
23 : ZIP: 07054
24 :
25 : COMPUTER READABLE FORM:
26 : MEDIUM TYPE: Floppy disk
27 : COMPUTER: IBM PC compatible
28 : OPERATING SYSTEM: PC-DOS/MS-DOS
29 : SOFTWARE: PatentIn Release #1.0. Version #1.30
30 : CURRENT APPLICATION DATA:
31 : APPLICATION NUMBER: US/09/749,340
32 : FILING DATE:
33 : CLASSIFICATION:
34 : PRIOR APPLICATION DATA:
35 : APPLICATION NUMBER: 09/067,676
36 : FILING DATE:
37 : ATTORNEY/AGENT INFORMATION:
38 : NAME: Renda, Barbara L.
39 : REGISTRATION NUMBER: 27,626
40 : REFERENCE/DOCKET NUMBER: 33,333-01
41 : TELECOMMUNICATION INFORMATION:
42 : TELEPHONE: 973-683-2153
43 : TELEFAX: 973-683-4109
44 : INFORMATION FOR SEQ ID NO: 6:
45 : SEQUENCE CHARACTERISTICS:
46 : LENGTH: 1275 amino acids
47 : TYPE: amino acid
48 : STRANDEDNESS: single
49 : TOPOLOGY: linear
50 :

```



```

Db 293 EEMERARAREANAYDFIMEPPNKFNTLVGEKAGMSSGQKORIALAARLVNPKILLIDEA 352
Qy 604 TSLADPKRSEGVQALERAEGRTTYIAHRLSTIKTAHNIYLVNCKIAEGOTHELDVD 663
Db 353 TSLADSESKSAVOALAEKASKGRTTYIAHRLSTIRADLIYLVNCKIAEGOTHELDVD 412
Qy 664 RCGAYRKRIYEAOKINQEKADALEDAEDLTADIAKIKITASSASDLDGKPTTIDRTG 723
Db 413 KRCGLYSLVMSQDIKADDESMSTYSTERKTS--LPLHSVSKISDF-----IDK-- 462
Qy 724 THKVSASAILSKRPETTPKYSILMTLKEFVASFNPKEIPMLIGLVSVLGGAGGPOPTAV 783
Db 463 AESTOSKEIS-----LPEVSLKILK----LNKPEMPVVLGTSLASVANGVHPFSTI 512
Qy 784 LYAKAISTSLPESQYSKLRHDADEWLSMFEVVGIIQFTOSTNGAFAVCSERLIRRAR 843
Db 513 IFKAITTMEG--NNDKTTTKHDAEISMIFFVLIGVICFVSFYMQGLFYGAGELTMRRLR 570
Qy 844 STAFRTILRODIAFEDKENSSTGALTSPLSTEDKHLGSGVGLTGLTLMSTLGAIIIT 903
Db 571 HIAFKMALYODIAMFEDEKESSTGGLTIIAIDIAQIGATGSRIGVLTOMATNMGLSVII 630
Qy 904 ALAIGKMLALVCISVVPVLLACGFYFYMLOPQSRSKLAEGSANFACETASITRTVAS 963
Db 631 SFYIGWEMTFLISTIAVPLAVTGMETFAAMTFGANKOKELKHAKIATFALENITRTVS 690
Qy 964 LRERBWEIYHAQOLAQGTSLISVLRSSLYASSOALVFPCVALGFYWGTLGHHEX 1023
Db 691 LRERKAFEBMYEEMLOTOHRNTSKAKQIIGSCYAFSHAFYFAAGFRGAVLIQAGRM 750
Qy 1024 DIFRFVCESEILFGASQATVSPFAPDGMKAKANAEPFRLEDRKPOIDNMSEBEKLE 1083
Db 751 TEBGMETVFTALAYGAMALIGKTLVLAPEYSKAKSGAAHFLALEKKNIDSRQEBGKPPD 810
Qy 1084 TWEGEIEFRNHRYPTRBQPYLRGLDLTVKBPQYVALVGPSCGKSTTILLERFYDA 1143
Db 811 TEGNLEFEVSEFFYPCRPDVFILRLSLIERKTYAFAVSGCGKSTSVQQLQLYDP 870
Qy 1144 IAGSLVNDKDISKLINSTRSFLSVSOEPTLYOGTIKENILGIVEDDVPPEFLIKAC 1203
Db 871 VOGQVLPEDVDAKELVQMLRSQIALVPEPVLFCNSIAENIAYGNSRVPVLDIEKEA 930
Qy 1204 KANITDFIATSLPEGFNTVYSGKMLSGQKORIVALAARLBDPKILLIDEATSLDSE 1263
Db 931 NMANIHSEFLEGPEKNTQYGLKGAQLSGQKORIALAARLQKPKILLIDEATSLDND 990
Qy 1264 SERVOALDAARGRITTAVAHRLSTIQADVIYVFDGKIVESGTHSELVOKGRYYE 1323
Db 991 SERVOALDKARTGRITCLVTHRLSAIONADLIYVLHNGKIKEGTHOELLNRDIYFR 1050
Qy 1324 LVNIQSL 1330
Db 1051 LVNAQSV 1057

```

RESULT 14

```

US-09-882-694-11
; Sequence 11, Application US/09882694
; Publication No. US20030009782A1
; GENERAL INFORMATION:
; APPLICANT: Davick, Jon
; APPLICANT: Maddox, Joyce
; APPLICANT: Gilliam, Jacob
; APPLICANT: Folkerts, Otto
; APPLICANT: Craesta, Oswald R.
; TITLE OF INVENTION: Compositions and Methods for Fumonisin
; FILE REFERENCE: 35718/208255
; CURRENT APPLICATION NUMBER: US/09/882,694
; PRIOR APPLICATION NUMBER: 09/351,224
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 11

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1263
; TYPE: PRF
; ORGANISM: Exophiala spinifera
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 157
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-882-694-11

Query Match      28.9%; Score 1940; DB 9; Length 1263;
Best Local Similarity 36.8%; Pred. No. 5,6e-136;
Matches 490; Conservative 231; Mismatches 483; Indels 128; Gaps 30;

Qy 54 ADKHHPRSSNNVAVNEVDALIHLPEDEROVLKTQLEERKIVNS--PFGIMRATKM 112
Db 2 ADSEKRPKNQDSSESS-----SHPPE-----KETEGISDYLRIETRYADKY 44
Qy 113 DILIMVISTICAIAASTFORIMLYOISY-----DEFEDELTKNLYEVYVL 158
Db 45 DMTLVNIALICALGASASLPLMSIIFGSPFNKNPNYNSGDGSPEARKADVDHVFVYL 104
Qy 159 GIGEFVTVVSYTGFIYTGHAQKIREYILESLIQNIGYEPKLGAGEVTTITLDTNL 218
Db 105 FIKGFVLTYSATATITSAIRTRTTLRNVFLTECTLOEWHFQKQNGALATAYVTNGNR 164
Qy 219 IODGISEKGLTTLATATEVTAFTIAYVKYWKALICSSFTYVALVLTMGSGSQFIITKYSK 278
Db 165 IOTGIAKEKLVTVQALSMFSAFVVALASOMKALITMTSTPIFLVTGICIAIDAQEA 224
Qy 279 KSLDSGAGGTVAEEVYSSIRNATAGTODKLAQYEVHLEDEAKWGTNRQIYMGFMIGA 338
Db 225 RITRIYRAVLAIEVYLSIRTVAFYAOQKWEKYDVFLQAHQGGKKSPNNGVLEST 284
Qy 339 MEGIAMTNGTGLGFMGSRFLVDGAY--DVGDILTVLMAIILGSSLSLGNVSNMAQFTNAYA 397
Db 285 EYFCITAAIALAFWKFRMTQNGEVADVGVTFVVLSTVLAATISMLAPSGSVVYQRI 344
Qy 398 AAKIFGTIDROSPLDPYSNEGKTLDFEGHIELRNVKHIYPRPVTWEDVLSMPAG 457
Db 345 FGSELFSLIDKPTQDPLDPSGQKQPEGLQGLEIQMLAFYRPRPAQVLRDNLITPAG 404
Qy 458 KTTALVPGSSGSGSYVGLVERFYMPKRGTVLLDGDIDKLNRLRQOISVSOEPLV 517
Db 405 KTTALVAGASGSGKSTVGLERWYLPSSGRILLDGLGELGQYKWKVRSRIRLVQOQBPVL 464
Qy 518 GTTIYKIRHGLGTYEENESDEKVELTENAKMANHDEFALTALPEGETVNGOGFILL 577
Db 465 RGTIFPNIANGFDEOROLPRE--KOMELVQKACKASNGDVFIELNPGETEVGERAGAL 523
Qy 578 SGGOKORIALAARVSDPKILLIDEATSLDTRKSEGVQALERAEGRTTYIAHRLST 637
Db 524 SGGORORIALAARSIISDPKILLIDEATSLDPRAEKVEQALNRVSKDITVIANKLAT 583
Qy 638 IKTANHIVLVNCKIAEGOTHELDVDRGAYRKLYEAQRI--NEQKADALEDAEDLT 695
Db 584 VYSAGNIAVISOGLIVEQTHHELIFEGCHYALVRAQDLGADDEQEHNE----- 632
Qy 696 NMDIAKIKTA--SSASDLDGKPTTIDRTGTHKSVSASAILSKRP--PETTKYSILMTLKRV 753
Db 633 -----KTLHEKAREAAAGERPALERTHTYTASQGDLEKRVPGTGLYSILKCI-LI 684
Qy 754 ASFNRPETIPY--MLIGLVSVLAGGCGPTOAVLYAKAISTSLPESQYSKLRHDADEWSIM 812
Db 685 MEYEOKNLWCPFLSTIYVILICATPPGALLFSRLTYFEL---SGHAAQERADPTIIM 741
Qy 813 FFYVGIQFTIYSTNGAFAVCSERLIRARSTAFETTLRODIAFEDKENSSTGALTSLP 872
Db 742 FFYVALAGNLVGYFTIGTWCNISQVYTHRYQAMFQRVLDQDIELDLIDPEQISALTSQL 801
Qy 873 STETKLSGVSQVLTGLTILMTSTLGAIIITIALAIGKMLALVCISVVPVLLACG----- 926

```



```

Db 802 SALPTOLQEL-----ISANPLIYVVGQHLQEC-STTSLMTEGPGGCW 846
QY 927 ---FYREYMAQFOSRS-----KLAVEGSANFACEATS---SIRTVASLTRERDWEIYH 975
Db 847 CTSTPAFGWLPQNSRDSEARSLGKRCRCMACRBSYDPPRLIDRSRPPSYLIGH 906
QY 976 AQIDAGCRISLISLVANSSLL-----YASSQALVFVCVALGF-----W-YGCTLLGHHHE 1022
Db 907 VE---QG---LAKIOSFMFGRCFGLHSQSMFLAIALGFCILAVDNWMLAVSTQLN--- 957
QY 1023 YDIFREFVCFS-EILFGASQAGTVESFAPDMGKAKNAAEFRRLFDKRPQIDNMSEGEK 1081
Db 958 -----FISSMAFCLPVCQAQIYLAFTSTPAKRSANNTLMRLTKPIRE-TEENKK 1010
QY 1082 LETVEG-EIEFRNVHRYFPPEOPVLRGDLTVKPGQYVALVGPSCGKSTTIALERF 1140
Db 1011 KGVVGGCPVDLEDIEFRYRQDSARVLRGVSMTIEPGFAVYAGASCGSTLALSERF 1070
QY 1141 YDAIAGSILVDGDISKLNINSTRSFLSVSOEPTLYOGTIKENILIGIVEDDVEEFLI 1200
Db 1071 YDPTSGRISFAHNIEMSPRLYRGHMSILVQOEPTLYQGSVREWVIAL-EAEISEE-LC 1128
QY 1201 KAKKDANIYFIMS--LPEGFNTVVSCKGMLSGGOKORVAIARALLRDKIILLDEATS 1258
Db 1129 QGRLPAPMLWISLSLYQAKRLAORGMQFSGGOKORIALAALLRNKLLILDATS 1188
QY 1259 ALDSESEKVVQAALDAARGRTTIAVAHRLSTIOKADVIYVFDGKIVESGTHSELVQOK 1318
Db 1189 ALDTQSERLVAQALDSESTRTTIAVAHRLSTIRNDVIYVFPANGRIAEIGTHAELOQR 1248
QY 1319 GRYEVLNLOSL 1330
Db 1249 GRYEMCLAQSL 1260

RESULT 15
US-09-873-409-2
; Sequence 2, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-409-2

```

Query Match 25.2%; Score 1690.5; DB 10; Length 812;
 Best Local Similarity 43.5%; Pred. No. 1,2e-117;
 Matches 366; Conservative 152; Mismatches 293; Indels 31; Gaps 7;

```

QY 489 LLDGHIKDLNLMRLQOISLVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKRANI 548
Db 1 MYDENDIRALNVRHRYDHIGVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKRANI 51
QY 549 AAKMANAHDEFTALPEGETNINOGRGFLSGGOKORAIARAVYSDPKILLDEATSALD 608
Db 52 AAREANAYDIMEFPKFNFTLVGEKQMGOKORAIARALRNPKILLDEATSALD 111
QY 609 TKSEGVQAALERAABGRRTTIVAHRLSTIKTANIVLVNGKIAEQTHDELVDRCGAY 668
Db 112 SESKSANQAALERAKGRRTTIVAHRLSTIRSDLIYTLKDMIAEKGAHAELMAKRGY 171
QY 669 KRLVEAQRIVEQKADALEDAEDLTNADIKIKTASSASDLOGKPTTIDRTGTHKSV 728

```

```

Db 172 YSLVMSODIKKADQEMESMTYSTERKTN--LPLHSVSIKSDF-----IDK--AEEST 221
QY 729 SSALISKRPPETTPKYSLMTLLKFAVSEFNREPEIRYMLIGLVSVALGGOPTOAVLAKA 788
Db 222 QSKRIS-----LPEVSLIKLK----LNKPEMFVVLGTLASVLNKTVMHVESITIRAKI 271
QY 789 ISLTLSESOYSKLRHDAEFWSLMFPVVGIIQFTTOSTNGAFAFVCSERLIRARSTAFR 848
Db 272 ITMGF--NNDKRTLLKHAETYSMIFVLGVICFVSYPMOGLFYGRAGEIILMRRLHAFK 329
QY 849 TIIRQDIAFDKENSNGALTSFSTETKHLISGVSYTLGITLMTSTTLGAAILTALAIG 908
Db 330 AMYQDIAWFDKENSNGALTSFSTETKHLISGVSYTLGITLMTSTTLGAAILTALAIG 389
QY 909 WKALVCISVYVPLVACGFREFYMLAQFOSRSKLAEGSANFACEATSIRTVASLIRER 968
Db 390 WEMFLLISTAPVLAVGMETTAAMTGFAKDKOLKHAGIATAEALENITIVSLTREK 449
QY 969 DWELIYHAQIDAQRTSLISVLRSSLLYASSQALVFVCVALGFYVGGTLGHHEDYIFRF 1028
Db 450 AFEQYEMLOQTOHRNTSKNAQIIGSCYAFSHAFYFAVYAGFAGYLLIQAGHMTPEGM 509
QY 1029 FVCSSEILFQAQSGITVFSFAPDMGKAKNAAEFRRLFDKRPQIDNMSEGEKLETVGE 1088
Db 510 FIYFTALAYGAMAIGKTLVLAPEYSKAKSGAHLFALLEKKNPNTDSRSQEBKKPDTECN 569
QY 1089 IEFENVHFRYPTREOPVLRGDLTVKPGQYVALVGPSCGKSTTIALLERFYDAIGSI 1148
Db 570 LEFEVSEFTPCRPDVTILKGLSITERGKIVAVGSSGCKSTSVQDLQKLPVQGV 629
QY 1149 LVGDKDISKLNINSTRSFLSVSOEPTLYOGTIKENILIGIVEDDVEEFLIKACKRANI 1208
Db 630 LFDGVDAKELNVQKLRSGQIAIVPEVLEFNCISIAENIAYGDNVSVVPLDEIKKANANANI 689
QY 1209 YDFIMSLEPGFNIVYVSGKGMLSGGOKORVAIARALLRDKIILLDEATSALDSESEKV 1268
Db 690 HSFTEGLPEKYNQVGLKQALSGGOKORLAIRALLQPKIILLDEATSALDSESEKV 749
QY 1269 QALDAARAGRTTIAVAHRLSTIOKADVIYVFDGKIVESGTHSELVQOKGRYELVNIQ 1328
Db 750 QHALDKARTGRTCLVYTHRLSAIONADLIYVHLNGKIKEGSTHOLLNRNDIYTKLVNAQ 809
QY 1329 SL 1330
Db 810 SV 811

```

Search completed: April 1, 2003, 16:08:22
 Job time : 34 secs

XX The present sequence represents a multiple drug resistance protein, designated atdD, which is isolated from *Aspergillus nidulans*. The atdD polynucleotide is used to produce nucleic acid molecules which are useful in diagnostic assays for the detection of atdD polynucleotides. The atdD polypeptide may be used to identify agents which inhibit multiple drug resistance. These agents may be useful for the treatment of fungal diseases. New antifungal compounds may also be identified using the gene disruption or gene replacement strains of *Aspergillus nidulans*.

XX Sequence 1334 AA:

Query Match 100.0%; Score 6721; DB 20; Length 1334;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MSPLFNPSPETAMREPASTSTTEQASTPHAADEKILISLDSAPSSTATPADKEHRP 60
DB 1 MSPLFNPSPETAMREPASTSTTEQASTPHAADEKILISLDSAPSSTATPADKEHRP 60
QY 61 KSSSSNNAVSVNEVDALIAHLPEDERQVLKQLEEKVNISEFGIMRWATKMDILMTVS 120
DB 61 KSSSSNNAVSVNEVDALIAHLPEDERQVLKQLEEKVNISEFGIMRWATKMDILMTVS 120
QY 121 TICAATAASTFORIMLYQISTDEFDLTKNLYFYLLGIGERYVYVSTVGFTYTGERRA 180
DB 121 TICAATAASTFORIMLYQISTDEFDLTKNLYFYLLGIGERYVYVSTVGFTYTGERRA 180
QY 122 TICAATAASTFORIMLYQISTDEFDLTKNLYFYLLGIGERYVYVSTVGFTYTGERRA 180
DB 122 TICAATAASTFORIMLYQISTDEFDLTKNLYFYLLGIGERYVYVSTVGFTYTGERRA 180
QY 181 TOKIREYTESILRONIGFEDKIGAGEVTTTITADTNLIQDGISEKVGTLTALATFVRA 240
DB 181 TOKIREYTESILRONIGFEDKIGAGEVTTTITADTNLIQDGISEKVGTLTALATFVRA 240
QY 241 FIIAVYKWKALICSSITVALVLMGGSGQFLIKYSKSLDSYAGAGTVABEVISIRN 300
DB 241 FIIAVYKWKALICSSITVALVLMGGSGQFLIKYSKSLDSYAGAGTVABEVISIRN 300
QY 301 AFAFGTQDIAQOYEVHDEAEKNGKNOIYMGFMIGAMFGLMYSNYGLGFWMGSRFLVD 360
DB 301 AFAFGTQDIAQOYEVHDEAEKNGKNOIYMGFMIGAMFGLMYSNYGLGFWMGSRFLVD 360
QY 361 GAVDVGDIITVLMALIGSFGISGNVSPNAQFTNNAVAATAKIFGTIDRQSPIDPYSNECK 420
DB 361 GAVDVGDIITVLMALIGSFGISGNVSPNAQFTNNAVAATAKIFGTIDRQSPIDPYSNECK 420
QY 421 TLDHFEGHIELRNKHIYPSREPVYMEDVSLMPAGKTTALVGPSSGSKSTVGLVERF 480
DB 421 TLDHFEGHIELRNKHIYPSREPVYMEDVSLMPAGKTTALVGPSSGSKSTVGLVERF 480
QY 481 YMPVGGTVLLDGHDKDLRWLRQOISVSOEPLYFGTTIKNTRHGLIGTRYENESD 540
DB 481 YMPVGGTVLLDGHDKDLRWLRQOISVSOEPLYFGTTIKNTRHGLIGTRYENESD 540
QY 541 KYRELLENAAKANAHDFITALPEGETVNGQGFLLSGGQKRIAIARAVVSDPKILL 600
DB 541 KYRELLENAAKANAHDFITALPEGETVNGQGFLLSGGQKRIAIARAVVSDPKILL 600
QY 601 DEATSDLTGKSGVVOALERAAGEGTTIVIAHRUSTIKTANHIVLVNGKIAEGCTHDE 660
DB 601 DEATSDLTGKSGVVOALERAAGEGTTIVIAHRUSTIKTANHIVLVNGKIAEGCTHDE 660
QY 661 LYDRGGAARKLVLEAQRINQKREADALEDDADLTNADIAKIKTASASSDLDGKPTTID 720
DB 661 LYDRGGAARKLVLEAQRINQKREADALEDDADLTNADIAKIKTASASSDLDGKPTTID 720
QY 721 RTGTHKSVSSAITSKRPPETPKYSIMTLTKFVASFNREPIYMLIGLVSYLAGGGQPT 780
DB 721 RTGTHKSVSSAITSKRPPETPKYSIMTLTKFVASFNREPIYMLIGLVSYLAGGGQPT 780
QY 781 QAVLYAKAISTLSLPESQYSLRHADAFWLSMFVVGIIQFTQSTNGAFAVCSERLIR 840
DB 781 QAVLYAKAISTLSLPESQYSLRHADAFWLSMFVVGIIQFTQSTNGAFAVCSERLIR 840

```

```

QY 841 RARSTAFRTILNODIAFPDKENSTGALNSFLSTENKHLISGSGVTLLGTLMTSTTLGA 900
DB 841 RARSTAFRTILNODIAFPDKENSTGALNSFLSTENKHLISGSGVTLLGTLMTSTTLGA 900
QY 901 IITAIAGKIALVLCISVVPVLACGFYRFYMAQFQSRKATYAGSANFACATSTIRT 960
DB 901 IITAIAGKIALVLCISVVPVLACGFYRFYMAQFQSRKATYAGSANFACATSTIRT 960
QY 961 VASLFRERDWEIYHAQDLAGRTSLISVLRSSLLYASSQALVFQVALGFVYGGTLLGH 1020
DB 961 VASLFRERDWEIYHAQDLAGRTSLISVLRSSLLYASSQALVFQVALGFVYGGTLLGH 1020
QY 1021 HEYDIFRFVFCSEILIFGQASGTFSPFAPDMGKAKNAAEFRRLPDRKPOIDNNSEBE 1080
DB 1021 HEYDIFRFVFCSEILIFGQASGTFSPFAPDMGKAKNAAEFRRLPDRKPOIDNNSEBE 1080
QY 1081 KLETVGEIEFRNVHFRYTRPEQPYLRGLDLTVRPGQVVALVGPSSGCKSTTIALLERF 1140
DB 1081 KLETVGEIEFRNVHFRYTRPEQPYLRGLDLTVRPGQVVALVGPSSGCKSTTIALLERF 1140
QY 1141 YDAIAGSILVDKDISKLNINSYRSFLSVSQEPTLYOGTIKENILLAGIVEDDVPEEFLI 1200
DB 1141 YDAIAGSILVDKDISKLNINSYRSFLSVSQEPTLYOGTIKENILLAGIVEDDVPEEFLI 1200
QY 1201 KACKQANITDFIMSLPEGRNTVYSGKGMLSGGQKQRAVIAAALLRDPKILLDEATSL 1260
DB 1201 KACKQANITDFIMSLPEGRNTVYSGKGMLSGGQKQRAVIAAALLRDPKILLDEATSL 1260
QY 1261 DSESEKVOALDAARAGRTTIAVAHRLSTIQADVIYFDPQKIVESGTHSELVQKKR 1320
DB 1261 DSESEKVOALDAARAGRTTIAVAHRLSTIQADVIYFDPQKIVESGTHSELVQKKR 1320
QY 1321 YVELVNLQSLGKH 1334
DB 1321 YVELVNLQSLGKH 1334

```

RESULT 2

AA16434 AAY16434 standard; Protein; 1349 AA.

AC AAY16434;

DT 10-AUG-1999 (first entry)

DE Multiple drug resistance protein of *A. fumigatus*.

KW Multiple drug resistance protein; AfUMDR1;

KW multiple drug inhibition activity.

OS *Aspergillus fumigatus*.

PN US5914246-A.

PD 22-JUN-1999.

PF 08-MAR-1996; 96US-0612734.

PR 08-MAR-1996; 96US-0612734.

PA (ELIL) LILLY & CO ELI.

PI Peery RB, Skatrud PL, Tobin MB;

DR WPI; 1999-370499/31.

PS N-PSDB; AAX60201, AAX60202.

PT Drug resistance gene from *Aspergillus fumigatus* useful for

PS determining fungal multiple resistance inhibition activity

XX Claim 3; Columns 21-28; 25pp; English.

CC The present sequence represents the multiple drug resistance protein of

CC Aspergillus fumigatus, designated AfumDRI. The specification describes
CC a method whereby cultured transformed fungal cells are used to
CC determine the multiple drug inhibition activity of a compound.
CC The method is useful for providing nucleic acid molecules which
CC encode multiple drug resistance proteins from Aspergillus fumigatus
CC (Afuf) which are useful for determining the MDR inhibition of a compound.
XX

Sequence 1349 AA:

Query Match 80.5%; Score 5407.5; DB 20; Length 1349;

Best Local Similarity 79.1%; Pred. No. 0;
Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;

QY 5 EFNPLSPETAMREPASTTEBQASPHAADEKKILSDLSAPSSATPADKREKSS 64
DB 5 EFGASSREKSL-EDLQVATLEKGRSTSSGADNEKPHDHSLSDTIMAPPD----GKKRD 59
QY 65 SNNASVNEVDALIAHLPEDEROYLKQLEIKVINISFEGLMKATKMDILINVISITCA 124
DB 60 HGRAVDLND-DLFAHLOHEKEEVLKROLDAPEVKVSFTLYRYASRKDILILVSAICA 118
QY 125 IAA-----ASTFORIMLOISYDEYDELFRNVLVEFYLGIGEEVTVYST 170
DB 119 IAAGAALPLFTILFGSLASFQGISLCTMPEHEFKLTKNVLVEFYLGIAEVTVYST 178
QY 171 VGEITYGEHATOKIREYVESILRONIGYFDKLAGEVTTTRITADINLIDGISEKVLG 230
DB 179 VGEITYGEHATOKIREYVESILRONIGYFDKLAGEVTTTRITADINLIDGISEKVLG 238
QY 231 LRALATPFAFIAYKYKMLALIGSSTIYALVLMGGSGQFIKYSKSLDSTYGAGTV 290
DB 239 LFAFAFVFAFIAYKYKMLALIGSSTIYALVLMGGSGQFIKYSKSLDSTYGAGTV 298
QY 291 AEEVISINAPAFGTDKAKOYEVEHLEAEKMGTKNOVMGFMIGAMGLMYSNYGLG 350
DB 299 AEEVISINAPAFGTDKAKOYEVEHLEAEKMGTKNOVMGFMIGAMGLMYSNYGLG 358
QY 351 FWMGSRFLVDGAVDGIILVMAILIGSFLGNSPNAQFTNAAVAAAKIGCTIDROS 410
DB 359 FWMGSRFLVDGAVDGIILVMAILIGSFLGNSPNAQFTNAAVAAAKIGCTIDROS 418
QY 411 PLDPYSNEGKTLDFEGHIELRNKHTYPSRPVYMEVDSLSMPAGKTALVGPSSGK 470
DB 419 PLDPYSNEGKTLDFEGHIELRNKHTYPSRPVYMEVDSLSMPAGKTALVGPSSGK 478
QY 471 STYVGLVEFPYMPRGTVLLDGHIDKLNLRMLROOISLVSQEPVLEGTITYNIRGLI 530
DB 479 STYVGLVEFPYMPRGTVLLDGHIDKLNLRMLROOISLVSQEPVLEGTITYNIRGLI 538
QY 531 GTFYENESDEKVELLENAKANAHADFTALPEGYETNVGQGFLLSGOKORAIARA 590
DB 539 GTFYENESDEKVELLENAKANAHADFTALPEGYETNVGQGFLLSGOKORAIARA 598
QY 591 VSDPKILLDEATSLDTRKSEGVVQALERAAGRTTIVIAHRLSTIKAHNIVLVNG 650
DB 599 VSDPKILLDEATSLDTRKSEGVVQALERAAGRTTIVIAHRLSTIKAHNIVLVNG 658
QY 651 KLABOGTHELVDRGAYRLEAORINBOKEADLE---DADADELTNADIKITASS 707
DB 659 KLABOGTHELVDRGAYRLEAORINBOKEADLE---DADADELTNADIKITASS 718
QY 708 ASDDLDC---KPTTIDRTGTHKSVSAILSKRPETTPKYSLMTLILKFAVSFNRPPIY 763
DB 719 SSNSLDAVEDEKALREKRGTKQSVSAVLSKVPQEFKYSLMTLVKFGANRNPBLGY 778
QY 764 MLIGLVSVLAGGQPTQAVLVAKAISTLSLPSQYSKLRHDDFMSLMFVVGIOFIT 823
DB 779 MLIGLVSVLAGGQPTQAVLVAKAISTLSLPSQYSKLRHDDFMSLMFVVGIOFIT 838
QY 824 QSTNGAFAVNCSERLIRRASTAFRTILRODIAFPDKEENSTGALISFSTETKHSQVS 883
DB 839 LSTNGAFAVNCSERLIRRASTAFRTILRODIAFPDKEENSTGALISFSTETKHSQVS 898

QY 884 GVTGLTILMTSTLGAIIITAIAGWKALVLCISVVPVLLACGRFFRYMLAQOSRSLA 943
DB 899 GVTGLTILMTSTLGAIIITAIAGWKALVLCISVVPVLLACGRFFRYMLAQOSRSLA 958
QY 944 YEGSANACATSSIRIVASTREBDVWELYNHOLDQGSTLSIVRSLSLTVASSQALV 1003
DB 959 YEGSANACATSSIRIVASTREBDVWELYNHOLDQGSTLSIVRSLSLTVASSQALV 1018
QY 1004 FFCVVALGFMYGGLTLGHEHEDIFRFVCFSEILFGAOSAGTVSEFADMDKAKNAAEFR 1063
DB 1019 FFCVVALGFMYGGLTLGHEHEDIFRFVCFSEILFGAOSAGTVSEFADMDKAKNAAEFR 1078
QY 1064 RLDFRKPQIDNWSSEKLETVEGELIEFRVNHRRYPRPQPVLRGLDLYKPGQYALV 1123
DB 1079 RLDFRKPQIDNWSSEKLETVEGELIEFRVNHRRYPRPQPVLRGLDLYKPGQYALV 1138
QY 1124 GPGCGKSTTIALLERDYALAGSILVDGDISKLNINSTRSFLSLVSOEPTLYQGTIKE 1183
DB 1139 GPGCGKSTTIALLERDYALAGSILVDGDISKLNINSTRSFLSLVSOEPTLYQGTIKE 1198
QY 1184 NILGIYEDVPEBEFLKACKDANIYDFIMSLEPGFNTVGSKGMLSGGOKORVALARA 1243
DB 1199 NILGIYEDVPEBEFLKACKDANIYDFIMSLEPGFNTVGSKGMLSGGOKORVALARA 1258
QY 1244 LRLDPKILLDEATSLDSESEKVVQALDAARGRTTIVIAHRLSTIKAHNIVLVNG 1303
DB 1259 LRLDPKILLDEATSLDSESEKVVQALDAARGRTTIVIAHRLSTIKAHNIVLVNG 1318
QY 1304 KIVSGTHSELVQKKGKRYELVNLQSLGKH 1334
DB 1319 KIVSGTHSELVQKKGKRYELVNLQSLGKH 1349

RESULT 3

AAB62495
ID AAB62495 standard; Protein; 1333 AA.

XX AAB62495;

XX 09-JUL-2001 (first entry)

DE P. chrysogenum ABC transporter a10 sequence.

KM beta-lactam; micro-organism; ABC transporter; Arp-binding cassette;

KM cephalosporin; a10.

XX Penicillium chrysogenum.

PN MO200132904-A2.

XX 10-MAY-2001.

XX 03-NOV-2000; 2000MO-EP11489.

PF 03-NOV-1999; 99EP-0203684.

XX 03-NOV-1999; 99EP-0203685.

PR 03-NOV-1999; 99EP-0203687.

PR 03-NOV-1999; 99EP-0203688.

PR 03-NOV-1999; 99EP-0203689.

PR 03-NOV-1999; 99EP-0203690.

PR 03-NOV-1999; 99EP-0203691.

PR 03-NOV-1999; 99EP-0203692.

PR 03-NOV-1999; 99EP-0203693.

PR 03-NOV-1999; 99EP-0203694.

PA (STAM) DSM NV.

PI Van Den Berg MA, Bovenberg RAJ, Driessen AJM, Konings WN;

PI Schuurs TA, Nieboer M, Westerlaken I;

XX WPI; 2001-291055/30.

DR N-PSDB; AAF83399, AAF83400.

xx Disclosure; page 22-27; 37pp; English.
 ps A new multiple drug resistance (MDR) protein (AAR9255), designated
 xx Afl-MDR1, of *Aspergillus flavus* was identified as the product of
 CC Afl-MDR1, of *Aspergillus flavus* was identified as the product of
 CC a cDNA clone isolated from a mycelial cDNA library of *A. flavus*.
 CC Afl-MDR1 can be synthesised by host cells, partic. *Saccharomyces*
 CC *cerevisiae*, transformed with vectors that provide for the expression
 CC of DNA encoding Afl-MDR1. Transformed host cells can be used to
 CC identify cpds. that inhibit fungal MDR activity.
 xx
 xx Sequence 1307 AA;
 Query Match 59.4%; Score 3992; DB 17; Length 1307;
 Best Local Similarity 60.2%; Pred. No. 0;
 Matches 790; Conservative 211; Mismatches 244; Indels 66; Gaps 11;
 60 PRSSSSNNNAVSEVDALIAHLPEDEROYLKTQLEIKVNISEFGIMRRATKMDILIMVI 119
 13 PKSPGTGSGT-----GHSVSHAEVLDRLHTPVSOIGFGIYRYATRWDAILFG 64
 120 SITCAIAA-----ASTFORIMLYOISDEFEYDELTKNVLYEVYLGIEFVT 165
 65 SALAALAGGALPLFTVLFGRLTSTFQDIATHTTYDHFHHELTKNVYFYLYGAAEVA 124
 166 VYVSTVGFYTGSHATOKIREYLESILNQNGIFPDKLAGEVYTRITADTNLIQGISSE 225
 125 IYLATVGFYTGSHVYQQRVEYFOALNQNTAFEDTLAGEITTRITADTNLIQGISSE 184
 226 KVGLTITATVYTAFTIYVYKWLALICSSITVALVLTMGSGSOFITIKSKSIDSYG 285
 185 KVGALTGTSTVYTAFTIYVYKWLALICSSITVALVLTMGSGSOFITIKSKSIDSYG 244
 286 AGTVAAEYISSIRNATAGTODKLAKEYVHDEAEKNGTKQIYMGFMIGAMFGLMYS 345
 245 RGSNAEDLIDSTRIYVAAENAOETLAKTESHDKABEGMKSKVIFALVGLICMYL 304
 346 NTGLGFMWGSREFLVDA--VDVGDILTVLMAILIGSFSGNVPNAQAFNNAVAAAAP 403
 305 NKGLGFMWGSREFLVESISNKGADVLTIMAAIILGSYNNGVAPNQALSDVAAAASKV 364
 404 GTIDRQSPIDPYSNECKTLDHFEGLHLENVKKIYPSREPVYMEDVSLSPKGTATV 463
 365 GTIDRQSPIDALSDQCKTEFVNGNIVLQNIKRVYPSREPVYVAHDLSCITPAKTTAFV 424
 464 GPSGSGSTVYGLVERFYVPVREGTVLDDGHDILMLRMLROOISLYSOEVPFGTTIK 523
 425 GPSGSGSTTISLERFYDPVAGTIMGDHDIOITLMLRMLROOISLYSOEVPFGTTIAE 484
 524 NTRHGLIGTYENESDEKAVRELEENAKMANAHDFITALPEGEYETVNGVGRGFLISGOK 583
 485 NTRGIGTSRFEKESYEIRKRREAAARMANHDFITALPEGEYETVNGVGRGFLISGOK 542
 584 RAIARAAYVSDPKILLDEATSAIDTKSEGVQALERAAGRTTYIAHRLSTIKTAHN 643
 543 RAIARAIIITKDPKILLDEATSAIDTKSEGVQALERAAGRTTYIAHRLSTIKTAHN 602
 644 IYVLVNGKTAEGTDEIVDRGAYRKIVAEQRINFEKADALDADAEDLTN----- 696
 603 IYVLVNGKTAEGTDEIVDRGAYRKIVAEQRINFEKADALDADAEDLTN----- 660
 697 -----ADIAIKTASASASDLGKPTTIDRTGTR-----KSVSAILSK-R 736
 661 MTEFFPKDYPGDESDIYSI--LSDASDI-----GLHTGEKQVRYSRMSLSHLM 708
 737 PRETTYKYSIMTLTKFVASFNRREIYMLIGLVSVLAGGGQTOAVLYAKAISTISLPE 796
 709 QPVEKAYEFMTLTKFLASNRREMPFLGLLGLASITLAGIOPSQAVLFFAKASTISLIP 768
 797 SOTSKLRHADFWSLMEFVVGIIQFIOSTNGAFAVCSRLRLRRASTAFRTILRODIA 856
 769 LEYPKLRHADFWSLMEFVVGIIQFIOSTNGAFAVCSRLRLRRASTAFRTILRODIA 828

QY 857 FPDKENSTGALSTSLSTETKHLSGVSGVLTGTLMTSTTLGAAIIAIALGKALVCI 916
 DB 829 FPDQENSTGALSTSLSTETKHLSGVSGVLTGTLMTSTTLGAAIIAIALGKALVCI 888
 QY 917 SVYVLLACGFYTFYMLAQOSKSLAYEGSANFACATISITVSLTREDWELIYA 976
 DB 889 SAVPALLMCFVVMWIERFORAKKAYQESASACBAASAIRTVSLTMTETALOSYQA 948
 QY 977 QLAOGRTSLISLRSSSLVASSOALVEFCVAGLFWVGTLGHHEVDIRFPVCESTL 1036
 DB 949 QLRRLQSLDPLTVKSSSLVASSOALVEFCVAGLFWVGTLGHHEVDIRFPVCESTL 1008
 QY 1037 FGAQSAQTVSFAPDMGKAKNAAEFRRLPQIDNWESEGEKLETVGEIEFRVNH 1096
 DB 1009 FGAQSAQTVSFAPDMGKAKNAAEFRRLPQIDNWESEGEKLETVGEIEFRVNH 1067
 QY 1097 RYTPRPQVPLRGIDLTVKQGYVALVPGSGCKSTTIALLEFYDAIAGSILYDGDIS 1156
 DB 1068 RYTPRPQVPLRGIDLTVKQGYVALVPGSGCKSTTIALLEFYDAIAGSILYDGDIS 1127
 QY 1157 KLININSRPSLSVSOEPTLYOGTIRENIIIGIYEDDVPREPLIKCKDANIYDFIMSLP 1216
 DB 1128 TLEMSSTRSHLALISQEPFLQGTIRENIIIGIYEDDVPREPLIKCKDANIYDFIMSLP 1187
 QY 1217 EGFNTVVGSKGMLSGGQKQVARIARALRDPKILLDEATSAIDSESEKVVQAAALDAA 1276
 DB 1188 EGFNTVVGSKGMLSGGQKQVARIARALRDPKILLDEATSAIDSESEKVVQAAALDAA 1247
 QY 1277 RGRITIAVHRLSTIQKADVIYFPDQKIVESGHSLSLVOKKRYEVLNLOS 1329
 DB 1248 RGRITIAVHRLSTIQKADVIYFPDQKIVESGHSLSLVOKKRYEVLNLOS 1300

RESULT 5

AAB62493
 ID AAB62493 standard; Protein; 1261 AA.
 XX
 AC AAB62493;
 XX
 DT 09-JUL-2001 (first entry)
 XX
 DE P. chrysogenum ABC transporter a7 sequence.
 XX
 KW Beta-lactam; micro-organism; ABC transporter; Amp-binding cassette;
 KW adenosine triphosphate-binding cassette; antibiotic; penicillin;
 KW cephalosporin; a7.
 XX
 OS Penicillium chrysogenum.
 XX
 PN WO200132904-A2.
 XX
 PD 10-MAY-2001.
 XX
 PE 03-NOV-2000; 2000MC-EPI1489.
 XX
 PR 03-NOV-1999; 99EP-0203684.
 PR 03-NOV-1999; 99EP-0203685.
 PR 03-NOV-1999; 99EP-0203687.
 PR 03-NOV-1999; 99EP-0203688.
 PR 03-NOV-1999; 99EP-0203689.
 PR 03-NOV-1999; 99EP-0203690.
 PR 03-NOV-1999; 99EP-0203691.
 PR 03-NOV-1999; 99EP-0203692.
 PR 03-NOV-1999; 99EP-0203693.
 PR 03-NOV-1999; 99EP-0203694.
 XX
 PA (STAM) DSM NV.
 XX
 PI Van Den Berg MA, Bovenberg RAJ, Driessen AJM, Konings WN;
 PI Schuurs TA, Nieboer M, Westerlaken I;
 XX WPI; 2001-291055/30.
 DR N-PDB; AAF83393; AAF83394.

XX Enhancing secretion of beta-lactam compounds from a micro-organism by
 PT enhancing adenosine triphosphate-binding cassette transporter activity,
 PT useful for producing e.g. penicillin and cephalosporins -
 XX
 PS Claim 8; Page 92-94; 116pp; English.
 XX
 CC The invention relates to a method for enhancing the secretion of beta-
 CC lactam compounds from a micro-organism that comprises enhancing ABC
 CC (adenosine triphosphate (ATP)-binding cassette) transporter activity of
 CC the micro-organism. The method is used for enhancing the production and
 CC secretions of beta-lactam antibiotics like penicillin and cephalosporins.
 CC Manipulation of the ABC-transport protein system provides a means for
 CC enhancing beta-lactam secretion. The present sequence represents the
 CC P. chrysogenium ABC transporter aa7 protein sequence.
 XX
 SO Sequence 1261 AA;
 Query Match 44.3%; Score 2978; DB 22; Length 1261;
 Best Local Similarity 47.3%; Pred. No. 8.5e-238;
 Matches 603; Conservative 244; Mismatches 383; Indels 44; Gaps 11;
 85 EROYVKTQLEIKVYISFEGLMRWATKMDILIMVSTICA-IAASTPQRIMLYQI--- 139
 6 EOEIIAQOITPIIKINLQLEFRTATKQWETIGVSEFLAATITAGAITTTPALLTGILIGS 65
 140 -----SYDEFYDELTKNLYFEVVLGIGFEVYVYVSVGFITGEHATOKIREYLE 190
 66 IQASWGGSPGDRNSSELTFRITFYVFLGELIVSCYIANTGIFTGIIISRIHERYLA 125
 191 SILRONITFPDKLAGEVTRITADTNLQDGSISEKVGILITLALTAFYAFIAYKTKWK 250
 126 ALLSNIAFPFNIGAGEISTRTTADANLIRGISEKVAACVAVVAAAFISFMRWR 185
 251 LALICSTIVLVLMGGSGQFIITYSKSLDSCAGGVAAEVSVIRNATAFGTODKL 310
 186 LTLIASSLICIAVFAAAGIMLTIRQOOWLGETAESGTIYEEVSSIRTVVGLNAQSEL 245
 311 AKOYEVLDEAEKMGTKNOIYVGMIGAMFGIMYSNYGIGFWMGSRFLVDAGVADYDILY 370
 246 VARYGSLAKKERFANNRNLISGALLGAVFAVITLAIIGFWMGSRFLVAGTSSVVDYLT 305
 371 VLMALISFSISGNVSPNAQAFNTNVAATAKTFITIDQ-SPLDPRYSDMGKTLDFEGHI 429
 306 IILAVTGTIACIGGIVPLOWFTTAAAGSRLYSTIDRRPSTANRFSBG-SLDSVLGHI 364
 430 ELRANKHLYPSREPTVWEDVSLMPACKTTALVGPSSGKSTVGVLERFEMPRGVTY 489
 365 ELQNVKHIYPSRPDIYVNDNLSDIEPGKTTAVGPSSGSKSTIELLERFIDPVSGLIL 424
 490 LDGHDIKDLNRLMRLQOISLVSQEPVLFCTTYKNIIRGLIGTKYENSEDKVRELIEVA 549
 425 LDGKHLSELSPMRLQOISLVQOQSPFLPATTFENIRYGLVGTCPENASRENIENIVGA 484
 550 AKMANAHPTALPREGYETNGQRFLLSGGOKRIATARAIVSPKILLDEATSAIDT 609
 485 TRLANAHDFTIKLPDGYDLVGEAGVLLSGGOKRIATARAISPRILLDEATSAIDT 544
 610 KSEGYYOALAEBAEGRTTYIAHRLSTIKTAHNIVLVNGKIAIOGTHDELVDGAYR 669
 545 MSESIVQAHIEKASQGRITTYIAHRLSTYKSAKTIIVLSGGOLIOGHNDFLQRLANGYS 604
 670 KLVEAQRINEQKEDALDADADLTNADIAKIKTASSASDLDKPTTIDRTG---TH 725
 605 RLAKQAANVLSER---KRPDDDSLGVPIPAKSVSESE-KKAPPLVSGGGLPFDIV 658
 726 KVSVAIIS-KRPETTYKYSILWTLKTVASFNRPEIPMYMLIGLVSVLAGGGOPTQAVL 784
 659 RKGAALASLRPPH---QYSQWTLIRFVNSHKDSIKLVSMGFASISQIGTAGAQAQAVF 715
 785 YAKAISTSLPESQYSKLRHADFWSLMFVVGIIQFITQISNGAFAFVCSRLIRRRAS 844
 716 LAKCLVALARPTESPOLRSKSETNLMAGHVVYLAFOFAYISAQASALGKCTERLIRLGD 775

QY 845 TAFRTILRODIAFPDKEENSTGALTSFLSTETKHLGVSQVTLGLTMTSTTLGAIIITA 904
 DB 776 LSFRLALFEDMSFPMEEHGVGALVSLFCTEPBAMAGMCVAINGYMALTFLLIGALIMTS 835
 QY 905 LAIGKRLALVCSVYPVILLACGFYRFYMLAQRQSKLAIYESANAFACETSITVAVSL 964
 DB 836 MAVGKRLGIVGAAVTPVLLMCGELFRFVMAQLEAHLROYTOETASLAEVASAITVLSL 895
 QY 965 TERPVWMEIYHQDQAQRTSLISVRSILYASSQALVFCVAGLGFYWGGLLGHHEVD 1024
 DB 896 NRESEVAVFHKLKAEQDSKISRSLSSGLFAFQSAPILCTALGLWYGGTLVSGEYG 955
 QY 1025 IFRFVCFSEILFGAQSAGTVSFAPDMGKAKNAAEFRRLFRRRPQIDMNSSEGEKLET 1084
 DB 956 LFOFILSFPAAVNICGDAAGSIRSSPDIAKAKLSTRLKGLLDROPQ-----RPLDF 1007
 QY 1085 VE-----GEIFRNVHRRYPRPQPVLRGLDLTVKRGQYVALVGPSCGCKSTIALLE 1138
 DB 1008 VEPATLGTTEFRNVHRSYPRPDQILKGLDLTVHKGYALVGPSCGCKSTIVALLE 1067
 QY 1139 REYDAIAGSILVDGKDLSKLTINSYRSLVSOBPTLYOGTKENILGIVEDDVPEEF 1198
 DB 1068 RYHPLAGIVTMDGLDLSMDMSAIRNOVALYDQETPLFQGTIRNLLGLDASKYSQBE 1127
 QY 1199 LKACKDANIYDFINSLEGEFTVYVSKGMLSGQOKORVALARALLRDPKILLDEATS 1258
 DB 1128 LEICKDANIILDFIRSLPLGFMTGCGKGNFSGOKRLAARALLRBPVLLDEVTS 1187
 QY 1259 ALDSESEKVOAALDAAARGRTTIVAHRLSTIQADYIYVVDQKIYESGHSLSVQOK 1318
 DB 1188 ALDSESSQVVAALDQAAKQRTTVAIAHRLSAVRNADILCFLEDVETIESGHAELIRRR 1247
 QY 1319 GRYEVLVNLQSLGK 1332
 DB 1248 GRFAMLSIQNIKK 1261
 RESULT 6
 AAM62871
 ID AAM62871 standard; Protein; 1408 AA.
 XX
 AC AAM62871;
 XX
 DT 29-SEP-1998 (first entry)
 XX
 DE Multiple drug resistance protein called CneMDR1.
 XX
 KW Multiple drug resistance protein; CneMDR1; Identification;
 XX antifungal compound; resistance.
 XX
 OS Cryptococcus neoformans.
 XX
 PN US5786463-A.
 XX
 PD 28-JUL-1998.
 XX
 PF 08-MAR-1996; 96US-0612521.
 XX
 PR 08-MAR-1996; 96US-0612521.
 XX
 PA (ELIL) LILLY & CO ELI.
 XX
 PI Peery RB, Skatrud PL, Thornewell SJ;
 XX
 DR WPL: 1998-436610/37.
 DR N-PSDB: AAV42347; AAV42348.
 XX
 PT Dna encoding multiple drug resistance protein of Cryptococcus
 PT neoformans - also related vectors and transformed host cells, and
 PT method for expression of the protein
 PS Claim 1; Columns 17-26; 19pp; English.

D	b	143	KIRKFFHAIINNOIGFVDHVDGELNTRLTDVSKINIBGKICGMFQOMAAFPFGGI	202
Q	y	243	IAYVKWKLALICSSITVALVTLWGGSGOPFIYKSKSDSYGAGGVAEVIISIRNAT	302
D	b	203	IGFRGKMLTLVLAISPVLGSLAGIWAIIISFDTKELOAYAKAGVAEVLAAIRVI	262
Q	y	303	AFGRQDLAKOYEVHIDAEKWKGNIVYFPAIGAFMGFMTNVLGFWMGSRFLVYDA	362
D	b	263	AFGQKKELEERYNNLEEARLGIKATVNMISMGAFLLIYASVLAWGYTSVLSKE	322
Q	y	363	VDVODITLMAALISFSGISGNVSPAQFPTNVAANAARJFGIDOSPDLVSMNGKTL	422
D	b	323	XSIOQVLTFFPAVLIGFSGASPNTEAPAAKGAAYEFINTIDKPSIDSESKNGIYF	382
Q	y	423	DHEGHTELRNKHIYPSREPYTMEVDVLSMBAGTKTALVGPSSGKSTVGVIEREYF	482
D	b	383	DNIGNLEFKNIHPSYSRDOYOLKGLNKVGSQGTVALVGNSSGKSTVQLORLD	442
Q	y	483	PVRGTVLDDHDKDLRLRQOISLVSOEPLVEGTTIYKNIHGLISTKYNESEDKY	542
D	b	443	PTEBVSIDSGDITRTIVRYRLREITIGVSOEPLFATTAENIRYG----RENYTMBE-	496
Q	y	543	RELLENAKAMANDFTTALPEGERNVGOGRGLVSGGOKORAIARAVSDPKILLDE	602
D	b	497	---IEKAVKENADVEIMKLPKHFEDVYBERGQSLSGGOKORAIARALVRNPKILLDE	553
Q	y	603	ATSLADTRSGEVOAALERAEGRTTVIAHRLSTIKTANIVLVANGKIAOGTHDELY	662
D	b	554	ATSLADTRSEAVOALDKARKREGRTTVIAHRLSTYRNADIIAGPFGYIVGQGHHELM	613
Q	y	663	DRCAGYKRLVBOIRINEQKPADLLEDADAEDLTNADIAMIKTASSASSDLDKPTTIDT	722
D	b	614	REKIYKRLVMTQAGNEIELGNEVGESEKNEIDNLMMS---SKDSASSLIRRRSTRRSTR	670
Q	y	723	GTHSVSASALISKRPPTTKYSILMTLTKVVASFNBPETPYMLIGVYSVLGGGQPIQA	782
D	b	671	GPHODKRLSTKEALDEBVDPPISFWRLK---LNSSEMPYFVYGJFCALVAGLOPAFS	726
Q	y	783	VLVAKAISTSLPESQYSKLRHDADPWSLMFFVVGJIOPTOSTNGAFAVCSERLIRRA	842
D	b	727	IIIFSKVGVFTRNMDDETK-RHDSNLFSLFLGLVISPITFFLPGFTFGKAGELLTRL	785
Q	y	843	RSTAFRTLLKODIAFPDKENSSTGALTSPLSTETKHLGSGVYTLGTTIMSTYGAAT	902
D	b	786	RYWYFKSMLKODVSWEPDNKNTGALTTRLANAGGVKATGARLAVIQTANJGTCII	845
Q	y	903	IATALGKLLVYCISVYVVLACGFRTVLAQFQSRKLAVEGSAFNACEATSSIRIYA	962
D	b	846	ISLTYGQTLTLTALAIPTAIAGVYEMKMLSGQALDKDELEGSKATXTEAIEFNRYV	905
Q	y	963	SLTRERDPMWEIYHAODAOGRSLISVLRSSSLYASSQALVEFCVALGFWYGGTLLGHE	1022
D	b	906	SLTRERQKFMWYQOSLIQIPRNMLKKAHPGIFRSTQAMMFYVLAACRFAPYIABEL	965
Q	y	1023	YDIFREFVCSSEILFGAQSAGIYFSFAPDMGCRKANNAAEFRLLFDKRPQIDNMSSEGERL	1083
D	b	966	MTEFNVLVFSALYFGLAMAAGVSSAPADYAKAKAVSASHIMIIEVPEIDSYSTGGLKP	1025
Q	y	1083	ETVGELEEFNNHPRVYTPREDOVLAGDLIVYKQGYVALVPSGCGKSTYIALLERYD	1142
D	b	1026	NTELEGNKEVEYFNFTPRDIDIVVLGLNILEYKKGTTALVYSSGGKSTVOLLERYFD	1085
Q	y	1143	AIASIIIVDKDISKLINYSRFLSIVSOEPLTYGTIKENLILGLIYEDVDPEEFLKA	1202
D	b	1086	PMAGTYFLDKRYNQLNVMQLRAHLGLIVSOEPLFPCSTAIENIAGDNSRVYSQEIETRA	1145
Q	y	1203	CKDANIYDTMSLPEGENTVYVSGKGLMSGGOKORAVAIARALLRDKPILLDEATSAIDS	1262
D	b	1146	AKRANIHQFTESLPDKRTNRYGDKGQLOSOGQORAIARALVRPHILLDEATSAIDT	1205
Q	y	1263	ESEVVOALDAARGRTTVIAVAHRLSTOKADYIVYFPOGKIVYSGHSELVOKKGRY	1322
D	b	1206	ESEVVOALDKARKREGTCVIAHRLSTIONMDLIYVINGGVKEGHTGQOLLAQKGYTF	1265


```

||||| :||:||||| :||: | :||||| :||| | :|||
Db 1084 ERFYPLAKVLLDKEIRLNVQMLRAHIGTSDPELFDCSIANIAYGNRSRVSOE 1143
OY 1198 FIKACKDANIYDFIMSLPEGEFTVVGSGKMGSGOKORAVAIARALNDPKLLIDEAT 1257
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1144 EIVRAKENIHAFIESLPNKSTKVGDKGTQSGOKRIARALVQPHILLIDEAT 1203
OY 1258 SALLDSESEKRVQAALDAAARGRTTAVAHRLSTIQKADVIYVFDGKIVESGTHSELVOK 1317
||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1204 SALLDSESEKRVQDALDKAREGFTCIYIAHRLSTIQADLIYVQNGRVEKHEGTHQDLAQ 1263
OY 1318 KGRYYELVLOS 1329
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 1264 KGIYFSMVSQA 1275

RESULT 9
AAW44073
ID AAW44073 standard: protein; 1280 AA.
AC AAW44073;
DE 26-JUN-1998 (first entry)
XX Human multidrug resistance P-glycoprotein; MDR1; prokaryotic homologue.
XX Homo sapiens.
XX Key Location/Qualifiers
XX FH Region 1..640
XX FT Region /note="MDR1-N from Fig 1"
XX FT Region 641..1280
XX FT Region /note="MDR1-C from Fig 1"
XX PN W09740160-A1.
XX PD 30-OCT-1997.
XX PF 24-APR-1997; 97WO-N000216.
XX PR 24-APR-1996; 96EP-0201094.
XX PA (UYGR-) RIJKSUNIV GRONINGEN.
XX PI Bolhuis H, Konings WN, Van Veen HW, Venema K;
XX WPI: 1997-535844/49.
XX PT Prokaryotic homologue of human multiple drug resistance protein -
XX used to screen for compounds that inhibit, or avoid, drug resistance
XX PS Claim 10; Fig 1; 35pp; English.
XX CC The present invention describes a recombinant or isolated nucleic acid
CC (1), derived from a prokaryotic gene, which encodes at least a specific
CC and/or functional part of a transporter protein (7P), or its
CC derivatives, which has functional and/or structural similarity with the
CC P-glycoprotein (PG) encoded by the human multidrug resistance
CC (MDR1) gene. The present sequence represents the human MDR1 protein,
CC derived from MDR1-N and MDR1-C as shown in the specification in
CC figure 1. (1) is used to express recombinant proteins; its fragments
CC are also useful as probes and primers for detection and amplification
CC of related DNA. The protein produced, or cells expressing them, are
CC used to determine if substances can inhibit, or avoid, MDR proteins,
CC and in a screening method for identifying compounds that inhibit
CC transport of cytotoxic substances from cells. Also, cells with a
CC transmembrane protein, especially where expressed from (1), can provide
CC (additional) MDR, particularly for use as a model system to study
CC mechanisms of action of PG.
XX Sequence 1280 AA;

```

```

Query Match 37.0%; Score 2489; DB 18; Length 1280;
Best Local Similarity 42.1%; Pred. No. 3.8e-197;
Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

OY 94 EELKVNISFFGIMRYATKMDILIMVISTICALNAASTFORIMLY----- 137
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 28 KKKPPTVSVFSMEFRYSNNMDKLYMVGTLAALIHGALPLMLVFEGEMTDFANAGNLED 87
138 -----QISDEFEY---DELTFKNVLXFYVLIGGEFTVYVSVGFIYGEHATQKI 184
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 88 LMSNTNRSDINDTGFENMLDEDMTRYAVYSGIGAGVLVAATVQVSWCLAGROIHFI 147
OY 185 REYTESILRONIGYRDKLAGEVTRITADNTLIDQISEKVGTLTALATVFATIA 244
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 148 RKQFPAIMRQELGEMFDVHDGELNRLDDVSKINEIGDKMGFQSMATFFTCFIYQ 207
OY 245 YKATKFLALICSTYALVLMGGSQFLIKYSKSLDSYGAGVYAEVVISRNATAF 304
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 208 FTRGMKLTIVIIAISPVGLSAVMAKILSFTDKELLAYAKAGVAEEVLAIKRTVIAF 267
OY 305 GTQDKLAKQYEVHIDEAEKMGTKNOIVMGFMIGAMGLMYNVGIFPMGSRFVNDGAND 364
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 268 GGOKKELEPRYNNKLEBAKIKIKATTANISGAFLITVASTALAFWGTTLVSGEYS 327
OY 365 VGDILTVMALILIGSFLGNVSPNAQFTNAAVAAAKIFGIDROSPLDPYNEGKTLDH 424
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 328 IGOVLVFPFSVLIGAFVSGASPSIEAFANARGAAEIKIINDKPSIDYSKSGKRPDN 387
OY 425 FECHIELRNVKTHIYSPRPVYTWEDVSLSPAKGKTALVPGSGSGKSVVGLVERPMV 484
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 388 IKGNIIEFRNVHSPYSRKVKIKLGLMLKVOSGQVTAALGNSGCKSTVQLMORLYDPT 447
OY 485 RGVLLDGDHIDKDLNRMRLROQISLVSOEPVFGFTTYINRHGLTGTVENESSEDKVNE 544
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 448 EGMVSVSDGDIRTINRFLEITIGVSOEPVLEFATTIAENIRYG-----RENTYDDE--- 499
OY 545 LIENAKMANAHDFITLPEGEYETNVGKGFLLSGGOKRIAIARAVSDPKILLIDEAT 604
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 500 -LEKAVKEANAVDFIMKLPKFDTLVGERGAQLSGGOKRIAIARALVNPRIILLIDEAT 558
OY 605 SALLDSESEKRVQAALERAEGRTTIVIAHRLSTIKTAHVIYLVNKGIAEGTHDELVDR 664
||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 559 SALLDSESEKRVQALDKARKGRTTIVIAHRLSTVRADVIYAGFDGVIEKENHDELME 618
OY 665 GGAYRKLVEAQRINQEKADALEDDADLCTNADIAKIKITASSASSDLDGKFTTIDRTCT 724
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 619 KGIYFLVIMQTAGNEVE---LENA-----ADESKSIIDALEKSSNDRSLIRKST 668
OY 725 HKVSASAILSKRP-----PETPKYSLMTLLKTFVASFNRPEIPMYLIGLVFSVLGGG 777
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 669 RRSVRSQAQDRKLSGKKEALDESIPVSPWRIMK---LNTWEMPFYVGVCAIINGGL 724
OY 778 OPTQAVLAKAISTLSLPSQYSKLRHDDFWMSLGFVVGIIQFTQSTNGAFAVCSER 837
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 725 OPTAFATISKIIIGVFRIDDPETK-RQNSNLSLFLAGIISFTFFLQGTFFGKAGBI 783
OY 838 LIRRASTPFTILRODIAFPDKENSTGALPSLSTERKHSVGSVGLGITLMTSTL 897
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 784 LKRLRLKMYFRBMLKQDVSWFPDPRKTTGALTTRLANDAQAQKGAIGSLAVITTONIANL 843
OY 898 GAIIITIALIGKMLALVCISVVPVLACGYRFRYMLAQOFSRKTLAYEGSANFACEATSS 957
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 844 GGIITISFIYGMQLFTLLAIVPIIAIAGVEMKMLSGCALDKKELBEGAKIATEAIEN 903
OY 958 ITFVASLTERBYWEIYNAQDLAOGRTSLISYRSLSLAASSQALVFCVALGMYGTL 1017
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 904 FETVYSLTQEQFEHETIAQSLOVPRYNSLRKAHIGITFISFOAMMYFSYACCFRGAFL 963
OY 1018 LGHHEHDIDRFVCFSEILFGAOSAGTVESFAPDMGKANAAEFRRRLDPRQIDNNS 1077
||:||||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 964 VAHKLSMFEDEVILVFSAVVFGMAVGOVSSFADYAKAKAISAHIMITEKTPLDISTYST 1023

```

OY	1078	E G E L T E V I E G E I R N A H F E P T R P E O P Y L R G D L I V K R G O V A L V P S G C K S T I L L	1137
Db	1074	E G L M P R L B E G N V T G E V E F N Y P I R P D L P V L O G I S L E Y K K Q I L A L V S G C C K S T V O L L	1083
OY	1138	E R F Y D A I A G S I L V D K D I S K L N I N S Y R S F S I L V S O E P T L Y O G T I K E N I L L G I V E D D V E E	1197
Db	1084	E R F E D P L A G K V L L D G K E I K R L N O M L R A H I G I V S O E P I L F D C S I A E N I A V G N S R V S O E	1143
OY	1198	F L I A C D A N I Y D P I M S L P G F M V N V G S K G M S G G O K R V A I R A L L M D P K I L L D E A T	1253
Db	1144	E I V A A A E A N I H A F I E S L P K I S T K V G D K O Q T L S G G O K I A L R A L V A N Q P H I L L D E A T	1203
OY	1258	S A L D S E S E K V V Q A A L D A A R G T T I A V A H R L S T I O R A D Y I Y F F O G K I V E S G T S H S E L Y O K	1317
Db	1204	S A L D T E S E K V Q E A L D K A R G C R C I V I A H R L S T I Q A N D L I V E Q N R V K E H O T H Q Q L L A Q	1263
OY	1318	K G R Y E L V N I O S	1329
Db	1264	K G I F S M V S V O A	1275

RESULT 10
AAV58186
ID AAV58186 standard; Protein; 1280 AA.

AC	AA58186;
XX	
DT	14-MAR-2000 (first entry)

Human wild-type multidrug resistance-1 (MDR-1) protein

KM Multidrug resistance; MDR-1; P-glycoprotein;
KM Membrane efflux pump; haematopoietic stem cell; transduction;
KM bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KM gene therapy; gene replacement; genetic defect; thalassemia;
KM Gaucher's disease; sickle cell anaemia; ex vivo expansion;
KM cytokine; wild-type.

OS Homo sapiens.

Key	Location/Qualifiers
EH	185
FT	/note="This residue is Val in a mutant MDR-1 (AAY58187)"
FT	

PN W09961589-A2.

PD 02-DEC-1999.

PF 27-MAY-1999; 99WO-US11825.

PR 28-MAY-1998; 98US-0086988.

PA (STUD-) ST JUDE CHILDREN'S RES HOSPITAL

PI Sorrentino B, Buntling K;

DR WPI; 2000-072615/06.

XX
33

PT Ex vivo expansion of hematopoietic stem cells transduced with a
PT sequence encoding human multidrug resistance-1, used for bone marrow
PT transplantation -

PS Claim 10; Page 71-79; 113pp; English.

CC This sequence represents human wild-type multidrug
CC resistance protein MDR-1. MDR-1 is a transmembrane efflux
CC pump, responsible for the export of drugs from certain
CC cells, particularly cancer cells. Wild-type MDR-1 shows
CC increased resistance to ecoposide and decreased resistance to vinca
CC alkaloids compared with a mutant form (AA158187) where the Gly at
CC position 185 is replaced by Val The invention relates to transducing

CC haematopoietic stem cells with nucleic acid encoding an MDR protein
CC and culturing the modified cells. The modified haematopoietic stem
CC cells are useful in bone marrow transplantation (to reconstitute
CC haematopoietic systems in patients who have undergone chemotherapy or
CC radiation therapy) and in *ex vivo* gene therapy of genetic defects in
CC cells derived from haematopoietic stem cells, e.g., thalassemia,
CC Gaucher's disease, sickle cell anaemia or leukaemia. The modified
CC cells can also be used to identify factors involved in regulating
CC proliferation and differentiation in haematopoietic stem cells.
CC Haematopoietic stem cells that express MDR-1 will be protected against
CC chemotherapeutic agents, so can be engrafted while the patient is
CC undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC provides sufficient cells to permit standard biochemical analysis.
CC Overexpression of MDR-1 allows cytokine-driven expansion of
CC haematopoietic stem cells by at least 10-fold compared with a maximum
CC of 4-fold in known procedures.

SQ Sequence 1280 AA;

Query Match	37.0%;	Score 2489;	DB 21;	Length 1280;
Best Local Similarity	42.1%;	Pred. No. 3.8e-197;		
Matches 535;	Conservative 229;	Mismatches 448;	Indels 60;	Gaps 9

QY	94	BEIIVN5FELMAYKARKMOLLIVNISTICAIAA5SFORMLY-----	137
Db	28	KEKRPYVSVMEXYSNMWLDKLYVWGTIAIITGAGLPLMMLVGMETDIFANAGLED	87
QY	138	-----Q1SYDEFY-----DELTKNVLYFVVLGICEFYVVVSTVGFYTGZHAOKI	184
Db	88	LM5NITRSDINDGFEPMNLEEDMTTRAYVYSSGIGAGVLAAYIQV5FWCLAA9ROTJHKI	147
QY	185	REYLESLKRONIGYFDFKLAGEVYTTTITADTNLIJOGISEKGLITLALATVYATIA	244
Db	148	RKQFHHIMROELGMEFVHDGELNLTPLTDV5KININGIDKJGMFQ5MATPFGYIVG	207
QY	245	YKFKMTALIC5STVIALVLTVMGG5QFIKYSKSD5GAGGTVAEEYISRNATP	304
Db	208	FTBRMKLTVIAL5LSPVLGL5ANWMAIL5SETDKELLATAKAGAVAEVLAARTYIAF	267
QY	305	GTQDPLAKOYEVLHDEAEKMGWTKNOIYMGFIMGAMFGLMYSNGLGFWM5RFLVDGAVD	364
QY	365	VGDILTVLMIILIG5FSIGVNSPMAQAFNTNVA5AAKIFGTIDBQ5PLD5Y5N8GKTLDH	424
Db	328	IGQVLYTFP5VLICAF5FGA5P5IEFANRGAAYELFKIIDKPSID5IS5GHKPDN	387
QY	425	FEGHIELNKNYKHIP5R5PEVYVM5D5L5MPACKTILV5P5G5G5K5TV5GLVEFYEMPY	484
Db	388	IKGHL5EPNRYH5F5P5R5K5EYKILKGLMLNKYQ5Q5TV5ALV5G5G5G5K5T5YQ5LMQRLDPT	447
QY	485	RGYVLLDGHDIKIDLNLKMLRQOISLV5Q5E5PVL5E5T5Y5YK5NIRHGL5GT5K5YEN5E5D5K5RE	544
Db	448	EGMYSV5GQD5IRT5N5V5F5REL5IG5V5O5E5P5L5F5AT5I5A5EN5RY5G-----RENYTMD5---	499
QY	545	LIE5NA5K5NA5HD5IT5PL5PEG5E5TN5G5OR5GL5L5G5G5K5OR5I5A5R5A5V5P5K5ILL5DEAT	604
Db	500	-IE5A5V5E5A5N5A5YD5IK5PL5K5H5FD5ILV5E5RG5Q5L55G5G5K5OR5I5A5R5A5L5A5R5N5P5K5ILL5DEAT	558
QY	605	SALDPT5EGV5Q5AL5ER5A5EG5RT5I5V5A5H5RL5STIK5I5H5N5I5V5L5V5N5G5K5IA5Q5T5H5DEL5YDR	664
Db	559	SALDTE5E5A5V5Q5AL5DR5AK5RG5T5I5V5A5H5RL5ST5YR5N5AD5Y5LAG5PD5GV5I5E5K5N5DEL5MKE	618
QY	665	GGAT5R5K5L5VE5A5OR5INE5Q5E5AD5LE5D5AD5ED5LT5N5D5IA5K5I5K5A5S5A5S5D5LD5G5P5T5ID5T5GT	724
Db	619	KG5IF5K5L5V5T5Q5T5AG5NE5V5-----LE5NA5-----AD5E5S5E5ID5AL5E5M5S5N5D5R5S5L5IR5K5R5T	668
QY	725	HK5V5S5A5IL5K5RP5-----PET5PK5SL5T5L5K5F5A5S5F5NR5PE5I5PY5L5IG5L5E5V5L5AG5G	777
Db	669	R5R5Y5R5G5Q5A5D5R5K5I5T5E5AL5DE5I5P5P55F5M5K5I5-----L5N5T5E5P5Y5V5V5G5P5C5A5I5N5G5L	724
QY	778	Q5T5Q5A5V5I5A5K5I5TL5L5P5E5O5Y5K5L5R5D5AD5F5W5L5M5F5V5V5I5Q5T5I5T5G5N5G5A5F5A5V5C5R	837


```

Db      844 GTGIIISFIYGMQLTLLLLAIIVPIAIAVGEEMKLSGALKDKKELEGAGKITEAIEN 903
QY      958 IRRVASTIRRDWEIYTHAOLDAGRTSLSVLRSSLLYSSQALVFCVLAIGWYIGCTL 1017
Db      904 FRRVSLTQOKKEHEHMAQSLQVPRNSLRKAHIFGITFSFTQAMVYSYAGCFRGAYL 963
QY      1018 LGHHEVDIFFPFCESEILGASAGTVFSFAPDMGRKAKNAAEFRRLFPKRPQIDMMSSE 1077
Db      964 VAKKMSFEVYLVFSNVYGVGAAGVSSFADYAKAKISAHTIWIETPLIDSYST 1023
QY      1078 EGKELTEGEIEIEFRNRYHFRPRPEQVLRGLDLTVKPGQYVALVGPSCGKSTTALL 1137
Db      1024 EGLMPRTLEGNTVEGEVFNYPTRPDIPVLOGLSLVKKGQTLALVSSGCGKSTVQOLL 1083
QY      1138 ERYDYAAGSLYDGDISTLNINSTRFSLYSQEPYLGQTKENILIGIVDDVPEE 1197
Db      1084 ERYDYPLAGVLLDGRKIKRLNQWLRHGLIVSQEPILDFCSIAENIAYGDNRRVYSOE 1143
QY      1198 FLIKACKDANIYDFIMSLPGEFNTVVGSKGMLSGCGOKORVAIRALLRDPKILLDEAT 1257
Db      1144 EYRAAKENAIHAFIESLPKRYSTKVGDGKTOLSGGOKHAIARALVROPHILLDEAT 1203
QY      1258 SALDSESEKVVQAALDAAGRTTIAVAHRLSTIQKADVIYVEDQKIVESGTHSELVOK 1317
Db      1204 SALDTESEKVVQEBALDKARREGRTCIYIAHRLSTIQNADLIYFONGVYKHEGTHQOLLAQ 1263
QY      1318 KGRYVELYNQOS 1329
Db      1264 KGIYFSMVSVQA 1275

```

RESULT 12

AAEI8964 standard; Protein; 1280 AA.

AAEI8964;

21-MAY-2002 (first entry)

Human wild type P-glycoprotein (wt P-gp).

Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer; haematopoietic cell; cytotoxicity; cytosolic; P-gp.

Homo sapiens.

WO200210205-A2.

07-FEB-2002.

01-AUG-2001; 2001MO-US24560.

01-AUG-2000; 2000US-222313P.

(UNII) UNIV ILLINOIS FOUND.

Ruth A, Roninson I;

WPI; 2002-206182/26.

New human P-glycoprotein mutants for treating cancer, comprises an ability to confer increased resistance to chemotherapeutic drugs relative to wild type P-glycoprotein or P-glycoprotein having a glycine to valine substitution

Claim 44; Fig 3; 47bp; English.

The present invention relates to an isolated mutant human P-glycoprotein molecule. The invention is used as vaccine. An antibody specific for the mutant P-glycoprotein is useful for determining whether a human tumour sample comprises tumour cells expressing mutant human P-glycoprotein. A recombinant expression construct comprising nucleic acid encoding mutant

P-glycoprotein is useful for administering an increased amount of a chemotherapeutic agent to an individual with cancer, by transducing ex vivo haematopoietic cells from the individual with the recombinant expression construct, reintroducing the transduced haematopoietic cells into the individual and administering an increased amount of a chemotherapeutic agent without consequent haematopoietic cytotoxicity. The mutant P-glycoprotein is useful for treating and diagnosing cancer and for screening compounds for the capacity to disrupt binding of cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux from cells expressing the mutant P-glycoprotein. The present sequence is human wild type P-glycoprotein (wt P-gp).

Sequence 1280 AA;

Query Match 37.0%; Score 2489; DB 23; Length 1280;
 Best Local Similarity 42.1%; Pred. No. 3.8e-197;
 Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

```

QY      94 EEIKVNISFFGLMRYATKMDILINVIPTCAIAASTFORIMLY----- 137
Db      28 KKKPTVSVFEMFRYSNMLDKLYMVGTLAIHIGAGLPLMLVGEWTDIFANAGNED 87
QY      138 -----QISDEFT---DELTKNLYFVYIGIEFTVYVSTVGFIYGEHATOKI 184
Db      88 LMSNITNRSNDINDGFEMLEEDMTRYAYVYSGIAGVLAAYIOVSFMCIAAGROIHKI 147
QY      185 REYLESTLRONIGYFPEKLGAGEVTRTADTNLIODSISKVGLTLATFTYATFIA 244
Db      148 RKQFFHAIHQEIEMFEDVHDVGLNTRLTLDVSKINISGIGIKIGFOSMAFFFGFTVG 207
QY      245 YKRYWKALICOSTVALVLTMGGSQFIIKYSKSLDSYGAGTVGAEEVSSIRNATAF 304
Db      208 FTRGMKTLIVLAISPYLGSAVMAKILTSFTDKELLAYAKAGVNAEVLAIARTYIAF 267
QY      305 GTQOKLAKOYEYHLDDEAKWGTNOIYMGEMTIGAMFGLMYNTYIGFWMGRFLVDGAVD 364
Db      268 GQOKKELEYRKNNEEARIGIKRAITANISIGAFLLIYASALAFYVGTLLVSGEYS 327
QY      365 VGDILTIVMALILIGSFSGVSNPAQFTNVAANAIFGTIDROSPLDPPSNEKTLDH 424
Db      328 IGVLTIVFSSVLIGFSGVQASPSIEAFANARGAAYELFKITIDNKPISIDTSKSGHKPDN 387
QY      425 FECHIELNRKHIYPSREVTMEDVSLSPRAGKTTALVPGSGSKTVGLVERFYMPV 484
Db      388 IKGNLEFRNVAHFSYPSRKEVKIKGLNLKYGSGQVALVVGSGGKSTVQQLMQGLYPT 447
QY      485 RGYVLIDGHDKDNLRLRQOISLVSEPVLEGTTIKYKTRHGLIGIKRYENESDKYRE 544
Db      448 EGVNSVSDQDRIINVRLEIGVSOEPLVFTTIAENIRYG-----RENVWDE--- 499
QY      545 LIENAKKANAHDFTALPEGEYTNVGOGRFLSGGOKORIAIRAVVSDPKILLDEAT 604
Db      500 -TEKAVKANAIVDTIMKLPKFDVLVGERGAQLSGQKORIAIRALVRNPKILLDEAT 558
QY      605 SALDTKSEGVQAALERAEGRTTIVIAHRLSTIKTANIVLVNGKIAEOGTHDELVDYR 664
Db      559 SALDTESEAVYQVALDKARKGRTTIVIAHRLSTVANAIVAGFDDGVYVEKGNDELKKE 618
QY      665 GGATRKLYEAOIRINEOKADALEDAEDLINAADAKIKITASASSDLDGKPTTIDRIGT 724
Db      619 KGIYFKLVMTQAGNEV--LENA-----ADSKSEIDALENSSSDSRSLRKST 668
QY      725 HKVSSAATLSKRP-----PETPKYSILWTLTKRVASFNRPEIRPYMLIGVFSVLGGG 777
Db      669 RRSVAGSOADRKLSTKALDESDIPVSVFWRMK-----LNLTEMPYFVVGFCALINGL 724
QY      778 QPTQAVIYAKAISTLSLSESOYSKLRHADFEWSLMFVVGIIQFTTQSGNAAPVCSER 837
Db      725 QPAFAIRSKILIGVTRIDDPETK--RQNSNLSFLIFLALIGISFTTFLOGTFFKAGEI 783
QY      838 LIRARSTAFITILRODAFEDKENSFGALTSFSTETKHLSGVGTLLGTLMTSTLL 897
Db      784 LTKRLRYVFRSMNLKQDVSWEDDPKNTTGALTTRILANDAAOVKAIGSRLAVITONIANL 843

```

QY 898 GAATTTAATGKALVCSVPLAACGFRYMLAQOFSKSLAYEGSANFACENTSS 957
 DB 844 GCGIISITFYQMOLITLLAIPIIAAGVEMKMSGQALDKKKELEGAGKATEAIEN 903
 QY 958 IRTVASLTRBDWEIYHQAQDAGRTSLISVLRSSILVSSQALVEFCVALFMYGGLT 1017
 DB 904 FRTVSLTDQKKEHMTAOSLOVPPYRNSLKKAHITGITTSFTQAMMTFSTAGCFRGAYL 963
 QY 1018 LGHHEYDIFRFVFCSEILFGAQSAGTVSFAPDMGKAKNAAEFRRLPDRKPDIDNMSE 1077
 DB 964 VAHKLSFEDVLLFSAVVEGAMAVGVSEFADYAKAKISAHIIIMIEKTELLIDISYST 1023
 QY 1078 EGKELTVEGEIEFRVNRHRTPRPEOPVIRGLDLYVKEQYVALVPGSCGKSTTIAL 1137
 DB 1024 EGMPTLTGKNTFGEVNVNPPRPDIPVLOGISLEVKKGQTLALVSSCGKSTVQQL 1083
 QY 1138 ERFYDAIASILVDGDISLNTNSYRFLSLVSOEPLTYOGTITILLIGYEDDVPPE 1197
 DB 1084 ERYDPLAGKVLIDGKEIKRLNQMLRAHLGIVSOEPILEDCSIAENIAYGDNRSRVVSOE 1143
 QY 1198 FLIKACKDANIYDFINSLEGEFNTVYVSGKGMISGQKQKQVATARALLRDPKILLDEAT 1257
 DB 1144 EYBRAKEMANIHFIESLEPKKSTYVKGDKTQLSGGQKQRIATARALVRPHILLDEAT 1203
 QY 1258 SALDSESEVYQALDPAARGRTTIVAHRLSTIQKADVIYVFDQKIVESGTHSELVOK 1317
 DB 1204 SALDTESEKYOALDKARGRTCIYIAHRLSTIQADLIYVFPONGRVKHEGTHOOLAQ 1263
 QY 1318 KGRYELVNLDS 1329
 DB 1264 KGIYFSMSVQA 1275

RESULT 13
 ABB07266
 ID ABB07266 standard; Protein; 1280 AA.
 XX ABB07266;
 XX 26-MAR-2002 (first entry)
 DE Human BCRP protein related seq id No. 2.
 XX Stem cell; ATP transport; protein; ATP-binding cassette; antiparkinsonian;
 KW hepatocytic; neurodegenerative; cytosolic; antianemic; muscular; BCRP;
 KW cardiac; gene therapy.
 XX Homo sapiens.
 OS WO200192877-A2.
 XX 06-DEC-2001.
 PD 30-MAY-2001; 2001WO-US17459.
 PF 31-MAY-2000; 2000US-0584586.
 PR 29-MAY-2001; 2001US-0866866.
 XX (STUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 PA Sorrentino B, Schuetz J;
 PI WPI; 2002-114368/15.
 DR N-PSDB; ABA94365.
 XX Identifying a stem cell, for treating e.g., muscular dystrophy,
 PT myocardial infarction, Parkinson's disease, or neurodegenerative
 PT disorders, comprises detecting the expression of an ATP transport
 protein (BCRP) by a cell -
 PS Disclosure; Page 55-58; 87pp; English.
 XX

CC The invention provides a method of identifying and/or isolating a stem
 CC cell that involves detecting the expression of an ATP transport protein
 CC containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
 CC comprising stem cells. The isolated stem cells may be used in the
 CC treatment of diseases such as muscular dystrophy, degenerative liver
 CC disorder, myocardial infarction, Parkinson's disease, degenerative
 CC disorders of the brain, and for tissue regeneration or replacement.
 CC Haematopoietic cells can be used in bone marrow transplants (e.g., for
 CC treatment of leukemia), and for ex vivo gene therapy for treating blood
 CC diseases such as sickle cell anemia and thalassemia. The stem cells can
 CC also be used as cell targets in gene therapy protocols. The present
 CC sequence represents a sequence related to the BCRP for which no relevant
 CC information has been provided in the specification.
 XX

Sequence 1280 AA:
 SQ
 Query Match 37.0%; Score 2489; DB 23; Length 1280;
 Best Local Similarity 42.1%; Pred. No. 3,8e-197;
 Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

QY 94 EIKVNIISFEGIMRVATKMDILIMVISTICAIAAATFORIMLY----- 137
 DB 28 KEKKPVYVSVMRYSNMIDKLMVYGTLLAIIHGAGLPLMLVPGEMDIFANAGNLEB 87
 QY 138 -----QISTDERY----DELTKNLYFYVLGIGEEFYTVYSTGVFIYGEHATOKI 184
 DB 88 LMSNITRSDINDTGFEMLEEDMTYAYVYSGIGAVLAAVYQVFCIAGRQIHKI 147
 QY 185 REYLLSILRONIGFYFDKAGCVETRTITADNTLIDJISEKGLTLTAAATVAFPIA 244
 DB 148 RKQFHAIKROELGMPDVHDVGLNRLDDVSKINDEIGDKTGMFPQSAWTFEFTGIVG 207
 QY 245 YVRYMKIALICSTIYALVLTMGSGSQFIITYKSKSLDSYGAGTAAEEVSIIRNATAF 304
 DB 208 FTRGKMLTLYTIALISPVLGISAAYMAKILSFDKELTAYAKGAAVEYLAIRVIYAF 267
 QY 305 GTQDKLAKOYEYHLDAAEKRGTKNQIYMGFMGAMFLKSNYGLFGMGRFLVDGAVD 364
 DB 268 GCGKKELEERYKNMLEAKRIGIKKATYANISIAAFLLIYASVALAFVYGTTLVLSGEYS 327
 QY 365 VGDILFVLAAILISFSLGVSFNAQAFNAAVAAKRTGTIDROSLDLPDYSMEGKTLDH 424
 DB 328 IGVYLVFFSVYLIGAFSVGASPSIEFANAKAAEYIFITIDNKSIDSYSYSGKRPDN 387
 QY 425 FECHIELRVNKHITPSPREVTYMEDVSLSPACKTALVPSGSGKSTVVGLEFERYMYP 484
 DB 388 IKGNLEFRNVHFYSYPSRKEVKILKGLNKVQSGQYVALVNSCGKSTVYQMLQRLYDP 447
 QY 485 RGTVLDDGHDIKDLNRLMLRQOISLYSOEPLYEGTITTYKNIRHGLIGTYENESDKVRE 544
 DB 448 EGMVSVSDGDIRTNVFLNEIIGVVSQEPVLEATTIAENIRYG-----RENTYMD- 499
 QY 545 LIENAKMANAHDFITALPGYETVNGOGEFLISGQKORIAIARAVSOPKILLDEAT 604
 DB 500 -TEKAYKEANAYFIKLPKPFYLVGERGAQSLSGQKORIAIARALVRPKILLDEAT 558
 QY 605 SALDTESEGVYQALERAAGRTTIVAHRLSTYKTAHNTLVYLNKIAEOGTHDELVDYR 664
 DB 559 SALDTESEAVYQALDKARGRTTIVAHRLSTVRNADVIYAGDDGVYIEKQGHDLME 618
 QY 665 GGAIRKLVENQKINEQKEDALEDAEDITNADIKITASSASSDLGKPTTIDRTGT 724
 DB 619 KGIYFKLVYMTQAGNEVE--LENA-----ADESKSEIDALEMSNDSRSLIKRST 668
 QY 725 HKSVSAILISKRP-----PETTPKYSIMTLTKFVASFNRPPIYMLGIVSVYLAGG 777
 DB 669 RRSVRSQAODRKLSTKEALDESIPVVSFWRIK-----INTLEMPFYVGVFCALINGSL 724
 QY 778 OPTQAVLYAKAISTLSLPSQYSKLRLHADFWMSIAFFVYGIIOFTIQSNGAFAVCSR 837
 DB 725 QPAFAIIFSKIIQVFTRIDDEPK-RQNSNLPBLLFLALAGIISITFIFFQGFPGAGET 783
 QY 838 LIRRASTAFRTILRODIAFPDKENSTGALTSFSTETKHLHSGVGVTLGTLMTSTTL 897


```

Db      784 LKRLVYVFRSMRLKQDVWFDPPKNTTGTALTRLANDAAQVGAIGSRLAVTTONTANL 843
Qy      898 GAAIIITAIAGMKLALVCISVYVPLLAGCFYRFYMLAQOFSKLAAYEGSANTACEATSS 957
Db      844 GCGIITISFYGMQTLITLAIYPIIALAGVVEKMLSGCALKKKELEGKATATEAIEN 903
Qy      958 IRTVASLTERRDWEIYTHAOLDAQRTSLISYLRSSILYASSSQALVFCVAGFYWGTL 1017
Db      904 FRTVVSALTQEOKEFHEMYAQSLQVYRNSLRKAHIFGITESFTQAMMYFSYAGCFREGAYL 963
Qy      1018 LGHHEIDIRFPYCESEILIFGAQSAQTVSPFADMGKAKNAAEFRRLPRKQIDNMSE 1077
Db      964 VAHKLMSPFDVLIVFSAVVEGAMAVGQVSSPADYAKAKISAAHIIIMILEKPLIDISYST 1023
Qy      1078 ESEKLETVGEIEFRVHRYPTRPPOVYLRGDLITVKPOYVALVPGSCGKSTTIAL 1137
Db      1024 EGIIMPYTLGENTYTFGEVYVNTYPRDIPVLOGLSLEVKKQOTIALVSSCGKSTVQQL 1083
Qy      1138 EREYDAIAGSILVGDISKLNINSYRSFLSVSOEPTLYOGTITKENITLGIYEDDVPPE 1197
Db      1084 EREYDPLAGVLLDGKEIKRLANVQMLRAHLGIYSQEPILFDCSIAENIAYGDNSRYVSOE 1143
Qy      1198 FLTKAKDANIYDFMSLPEGFNTVVGSGMHLSCGQKQVATARALRDPKILLDEAT 1257
Db      1144 EYRAKKEANIHAIESLPKYSTKVGDKGTOLSGQOKORIALARALVRQPHILLDEAT 1203
Qy      1258 SALDSEKVVQALDAAARFTTIAVAHRLSTOKADVYVFDGQIVSGHSELVOR 1317
Db      1204 SALDTSEKVVQALDKARBGRTCIYIAHRLSTIONADLIVQNGRVKRGHGHQILLAQ 1263
Qy      1318 KGRYELVNLQS 1329
Db      1264 KGIYFSMVSVOA 1275

```

RESULT 14

AAB81064

ID AAB81064

AAB81064; (first entry)

Cynomologous monkey P-glycoprotein variant 1.

Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;

Macaca fascicularis.

MO200123565-A1.

05-APR-2001.

28-SEP-2000; 2000MO-DS26592.

28-SEP-1999; 99US-0156921.

12-OCT-1999; 99US-0158618.

(GENT-) GENTEST CORP.

Stocker PJ, Steinel-Crespi DT, Crespi CL;

MPI: 2001-316136/33.

N-PSDB; AAF86127.

Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein (PGP) and homologous PGP polypeptides are useful for predicting bioavailability of compound and increasing PGP transporter activity in cell

Claim 9, Page 57-59; 84pp; English.

This invention relates to a polynucleotide sequence encoding a cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the PGP protein. PGP, also known as multidrug transporter, MDR1 is a member of the ABC transporter superfamily. The enzyme serves as an efflux pump exporting small molecules across the cell membrane. The invention includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence and protein, and also that of an allelic variant. The PGP polynucleotide sequence is useful for increasing PGP transporter activity in a cell. Antisense sequences of the cDNA are useful for inhibiting PGP transport activity in a mammalian cell. They may also be used for increasing the bioavailability of a drug. The present sequence represents the cynomologous monkey P-glycoprotein.

Sequence 1280 AA:

Query Match 37.0%; Score 2488; DB 22; Length 1280;
 Best Local Similarity 41.8%; Pred. No. 4,6e-197;
 Matches 532; Conservative 228; Mismatches 466; Indels 46; Gaps 5;

```

Qy      90 KTOLEIKVNIISFGLRATKMDILIMVISTCAIAASTFORIMLY----- 137
Db      24 KDKKERKPTVSYFSMERNYSDMLKLYMVGTLAAITHGAGPLMLVFPEDMDTFANAG 83
Qy      138 -----QISYDEYDELTKNLYLYEYLGIGEVTVYVSTVGFLYTGHA 180
Db      84 NLGDIGALLTNSSNITDTVPVMMLEEDMTRAYAYSSGIGAGVLAAYIQVSMCLAAGRQ 143
Qy      181 TOKIREVYLESILRONIGYVDKLAGEVYTRIRADTLIDDGISEKGLTLTALATVTA 240
Db      144 IHRIRQOFHAIHQEIGMDVDVHGELNTRLIDVYSKINEGIGDKIGMFPQSNATFTG 203
Qy      241 FIITAYVYKMLALICSSSTIALVLTIMGSGQOFITIKSKLSDSYAGAGTAAEYISSIRN 300
Db      204 FIYGFTRGWLTLVYLAISVYLGISAIVAKILSSFTDKELLAVAKGAAVEYLAIRT 263
Qy      301 ATAFGTODKLAKOYEVHLDAEKMGKTRKQIYMGFMIGAMFGLMYSNYGLGFMGSRFLVD 360
Db      264 VIFAGGOKELERKNKMLEAKRIGIKATANTISGAELIYASVALFWYGTTLVLS 323
Qy      361 GAVDVGDLVNLAILGSSPSLGNSPNAQFTNAVAANAARIPCTIDROSPIDPYSNEK 420
Db      324 KEYSIGVLTVPFVSVLGASVCGQSPSIEFAARCAAEIEFIIDNKSIDYSKSGH 383
Qy      421 TLDFEESHIELRNKHYPRPEYTVMEDVSLSPACKTTLVGPSSGKSTVGLVERF 480
Db      384 KPDNIKGNLEFRNHYEYPSRKEVYKILGLNLKQSQYVALVNSCGKSTYQLMQRL 443
Qy      481 YMPVGTVLLDGDIDKDLNLRWLROQISLVSOEYVLEFGTIIYKNIRHGLIGTKYENESD 540
Db      444 YDPTEGWSVDGODIRTNVRFLEIIGVYSQEPVLPAFTIAENIRYG-----RED 494
Qy      541 KYRELINAKMANAHDFITALPBGTYNTGQRFLLSGQOKORIALARALVRPKILL 600
Db      495 VTMDLEKAVKANAYDFINKLPQFPTLVGERAQISGQOKORIALARALVRPKILL 554
Qy      601 DEATSALDTKSEGVQAALERRAAGRTTIIYAHRLSTIKTAHNIVLVYNGKIAEOGTHDE 660
Db      555 DEATSALDTSEAVVQVALDKARKGRTTIIYAHRLSTIVRNADYAGGDDGVYKGNHDE 614
Qy      661 LVDRGAYVRKLVENQRIEKGKADALDADADLTNADIAKIKTASASSDLDGKPTTID 720
Db      615 LMKKEGIFYFLVMTQTAGNEIE--LENNADESKSEIDTLEMSSHDGSSILRRKSTRRS 671
Qy      721 RTGTHKSVSAILSKRPETTPKYSLTWTLKFAVSFRPELPYTLIGLVSVYLAGGQPT 780
Db      672 VRSGQGDRLSTKEALDESIPVSEWRIMK---LNLTPPVFVGVPCAIINGGLQPA 727
Qy      781 QAVLYAAIATSLSPESQYKLRHADAFWSLMPFVVGIIQFIOTSTGCAFAVACSELRIR 840
Db      728 FAVYFSKIIGIFTRNDAAERK-RQNSNLFSLLPLVLDIVSFITTFLOGFTFGKAGELITK 786
Qy      841 RARSTAPRTLLRODIAFDREKSNSTGALTSFLSTETKHLGVSQVTLGTLIMTSTLGA 900

```

```

Db 787 RLRVYFRSMRLQDVSWEDDPKNTGALTTRLANDAQAQGAIGSKRIATITONIANIGTG 846
Qy 901 IITAIAGKMLAVCISSVVPVLACGFYRFYMLAOFOSRSKLAYESGSAFACATSSIRF 960
Db 847 IISILYGNQLTLLALAIYPIIAIAGVEMKMLSGALADKKELEBAGKATPAIAINPRT 906
Qy 961 VASLRENDVWEIYHAQDADQGRSLISVLRSSLLYASSQALVFCVALGFMYGTLTGH 1020
Db 907 VVSLVQEQKEFHMVYDQSLQVPRNSLRKAHIFGIFSFQAMMYFSYAGCFRGAYLVAH 966
Qy 1021 HEYDIFRFVCSSEILFGASGAGTVFSFAPDMGKANAFAERRLDRKPQIDNWESEBE 1080
Db 967 SLMSFEDVLVFSAVFGMAVGVSSFPADYAKAIVSAHHIIMITEKPLIDSYSTEG 1026
Qy 1081 KLEVEGEIEFRNVRHFRPTREPOVPLRGDLTVKPGQVVALVGPSCGSKSTTALLERF 1140
Db 1027 KNTEGNTFNEVYFNPTRIDIPYLOGLSLEVKKQGLALVSSGCGSKSTVQGLENF 1086
Qy 1141 YAIAGSILVGDKDIKSLINSTRFSLVSGDEPPLVYOGTKEINILLGIVEDDPEEPLI 1200
Db 1087 YPLAGKVLLDGRKEIKQLVQWLRHMLGIVSQEPILFDCSISEINLAYGDNSTRVSGEEIV 1146
Qy 1201 KACKDANITDFMSLPEGFNTVSGKGMVSGGOKORVIAIARALLRDPKILLDEFTSL 1260
Db 1147 RAKKENIAFIESLPNKYSTVNGDKGTQSGOKORIAIARLVQPHILLIDESTSL 1206
Qy 1261 DSESEKVVQALDAAARGRTTIAVAHRLSTIQKADIVYFDQGIYESGSELYOKGR 1320
Db 1207 DRESEKVVQALDAAARGRTTIAVAHRLSTIQKADIVYFDQGIYESGSELYOKGR 1320
Qy 1321 YVELYNLSLGR 1332
Db 1267 YFSMVSVOGAK 1278

RESULT 15
ID AAE18972 standard; Protein: 1280 AA.
XX
AC AAE18972;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human P-gp mutant, MDR1-E643K/E647K/E652K (8-B4).
XX
KM Human; P-glycoprotein; vaccine; tumour; chemotherapeutic agent; cancer;
KN haematopoietic cell; cytotoxicity; cytostatic; P-gp; mutant; mutein.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 643 /note= "Wild type Glu substituted with Lys"
FT Misc-difference 647 /note= "Wild type Glu substituted with Lys"
FT Misc-difference 652 /note= "Wild type Glu substituted with Lys"
FT FT
XX
PN W0200210205-A2.
XX
PD 07-FEB-2002.
XX
PE 01-AUG-2001; 2001MO-US24560.
XX
PR 01-AUG-2000; 2000US-222313P.
XX
PA (UNIT ) UNIV ILLINOIS FOUND.
XX
PI Ruth A. Roninson I;
XX
DR WPI; 2002-206182/26.

```

```

XX
PT New human P-glycoprotein mutants for treating cancer, comprises an
PT ability to confer increased resistance to chemotherapeutic drugs
PT relative to wild type P-glycoprotein or P-glycoprotein having a glycine
PT to valine substitution
XX
PS Claim 42; Fig 3; 47pp; English.
XX
CC The present invention relates to an isolated mutant human P-glycoprotein
CC molecule. The invention is used as vaccine. An antibody specific for the
CC mutant P-glycoprotein is useful for determining whether a human tumour
CC sample comprises tumour cells expressing mutant human P-glycoprotein. A
CC recombinant expression construct comprising nucleic acid encoding mutant
CC P-glycoprotein is useful for administering an increased amount of a
CC chemotherapeutic agent to an individual with cancer, by transducing ex
CC vivo haematopoietic cells from the individual with the recombinant
CC expression construct, reintroducing the transduced haematopoietic cells
CC into the individual and administering an increased amount of a
CC chemotherapeutic agent without consequent haematopoietic cytotoxicity.
CC The mutant P-glycoprotein is useful for treating and diagnosing cancer
CC and for screening compounds for the capacity to disrupt binding of
CC cytotoxic drugs with the mutant P-glycoprotein or to disrupt drug efflux
CC from cells expressing the mutant P-glycoprotein. The present sequence is
CC human P-glycoprotein (P-gp) mutant, MDR1-E643K/E647K/E652K (8-B4)
CC (multidrug resistance).
XX
SQ Sequence 1280 AA;
XX
Query Match 37.0%; Score 2488; DB 23; Length 1280;
Best Local Similarity 42.1%; Pred. No. 4,6e-197;
Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;
Qy 94 EEIKVNIISFFGLMRYATKMDIIMVISTICAIAASTFORIMLY----- 137
Db 28 KKKKPTVSIFSMERYSNMWLDKLYMYGLTAAIIHGAELPMLVGEMLDIRANGLNED 87
Qy 138 -----QISYDEFY----DELTKNVLVYVYLIGIEFVYVSTGVFTYGERATOKI 184
Db 88 LMSNITNRSIDINDTGFENMLIEDMTRYAYVYSGIGVLAIVYQVSEFCLAAGROIHKI 147
Qy 185 REYLESTLRONIGYFDKAGEVTRITADNMLDQGSSEVGLTFLATFVAFITIA 244
Db 148 KQPFHAIKROELGMDVHDVDELNTRLDVDSKINEGIDKIGMFQSMATFFGFYIG 207
Qy 245 YVKYKMLALICSTIYALVLTMGGSQFIIRKYSKLSIDYGAGVAAEVISSINATAF 304
Db 208 FTRGKMLTIVLAIISPVGLSAVAKILISFTDELLAAGAAGVAEVLAAITVIAF 267
Qy 305 GTQDKIAQYEVHLDAAEKWGTRKQIVMGFMIGAMFGLMSYNGLGFWMGSRFLVDGAVD 364
Db 268 GGQKKELERYNNLEAKKIGIKKAITANISIGAAFLIYASVALAFWGTVLVSEGS 327
Qy 365 VGDILTVLMAIILIGSFGNVSFNPAQAFNNAVAAAKITGTDROSPIDPYSENEKTI 424
Db 328 IGVQLTVFVSILGAFVSQASPSIEAFANAGAAVEIKIIDNKPISDYSKSHKDPN 387
Qy 425 FEGHIELRNVKHIYPSREPTYMEDVSLMPAKTTALVGPSSGSKSTVGLYERFMPV 484
Db 388 IKGNIIEFRNVHSPYSRKEVKILKGLNLKVSGQVVALVGNCGSKSTVQLMQGLYDPT 447
Qy 485 RGTVLLDGHDIQDLNRLRQOISLVSQEPVLFGTIYKNIHGLIGIKYENESDKVRE 544
Db 448 EGMVSVDGODIRTIWRFLEIIGVSSPEVLFATVIAENIRYG-----REVNTDE--- 499
Qy 545 ILENAKMANANDFTALPEGETVNGRGFLTSSGOKORIAAAYVSDPKILLDEAT 604
Db 500 -TEKAVKEANAYDFTMKLPKFDTLVGERGAQLSGGOKORIAIARLVNPNPILLDEAT 558
Qy 605 SALDTRSEGVVQALERAAGRTTIVIAHRLSTIAHIVLVANGKIAQGHDELVD 664
Db 559 SALDTRSEGVVQALDAAARGRTTIVIAHRLSTVNAADVIAFGDGVYIEKGNHDLAKE 618
Qy 665 GGAYRKLVEAQRINQKREDALEDDADBDLVNADIATIKRTASSASSDLDGKPTTIDRTGT 724

```

Db 619 KGIFYELVTWQAGNRVK---LENA-----ADKSKSKIDALKMSSNDSSILIRKST 668
QY 725 HKSVSSAILSKRP-----PETTPKYSILWTLKFWASFNRPEIPYMLIGLIVFSVLAGG 777
Db 669 RRSVRSQOQDRKLSRKLEADESIPVYSFRIRK----LNLTEPPIFYVGVFCAILINGCL 724
QY 778 QPTQAVLVAKAISTLSLPSQYSKLRHADFWSLMEFFVVGIIQITQSTNGAFAVCSER 837
Db 725 QPFAFIIFSKLIIGVTRIDDEPK-RQNSNLFSLFLALGIIISFTTEFLQGFTFGKAGEL 783
QY 838 LIRRASTAFRTILRODIAFPDKEENSTGALJFSLSTERKHLGSGVYLTGTLMTSTTL 897
Db 784 LKRLRYMFRSMLRQDVSWFDDPKNTGTALTRLANDAAQVKAIGSLAVITONIANL 843
QY 898 GAIIITAIIGWTLAVCISVVPVLLACGFYRPMQAQFSRSKLAVESGANFACEATSS 957
Db 844 GYGIIISFTYGMQJLILLAIYPIIAIAGVVEKKMLSGALKDKKELEGAKIATEAIEN 903
QY 958 IRTVASLTRERDVWEIYHAQLDAQRTSLISVLRSSSLVASSQALVEFCVALGFYGGTL 1017
Db 904 FRTVVSILTOEQKEFHHYAQSLQVYRNSLKAHIFGITFSFTQAMMYFSYAGCFREGAYL 963
QY 1018 LGHHEVDIRFEVCFSEILFQAQSACTVSPADMGKAKANAAEFRRLEPDRKPOIDNMSE 1077
Db 964 VAKHIMSFEVDVLLVFSAYVFGAMAVGOVSSFADYAKAKISAHHIIMIEKTPIDISYST 1023
QY 1078 EGEKLEFVEGEIEFRNVHFPTRPEOPVLRGLDLYKPGQVAVLPGSGCGKSTTIAL 1137
Db 1024 EGLMPNTLEGNTVEGEVFNYPTRPDIPVLOGLSLEYKKGOTLALVSSCGSKSTVQTL 1083
QY 1138 EREYDAIAGSILVDGDISKLINISYRSLVLSVSOEPTLYOQTIKENILGIYEDVPEE 1197
Db 1084 EREYDPLAGVLLDGEIKRLANQWMLRAHLGIYSOEPILEDCSIAENIAYGDNRSRYSOE 1143
QY 1198 FLIKACKDANIYPIFMSLPEGENTVYSGKGMISGCGOKORAVATRALLPDKILLDEAT 1257
Db 1144 EIVRAKENANIHAFIESLPKYSTKVGDKTOLSGGOKRIAIARALVRQPHILLDEAT 1203
QY 1258 SALDSESEKVVQALDPAAGRTTIAVHRSLSTOKADVIVYFDGKIVESHSELYOK 1317
Db 1204 SALDSESEKVVQALDKARBGRTCIYIAHRLSTIQANADLIYFONGRNVKEGHQOQLAQ 1263
QY 1318 KGRYELVNLQS 1329
Db 1264 KGIFYFSWVSVOA 1275

Search completed: April 1, 2003, 16:00:39
Job time : 91 secs

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes the need for transparency and accountability in all financial dealings.

2. The second part of the document outlines the various methods and techniques used to collect and analyze data. It includes a detailed description of the experimental procedures and the statistical analysis performed.

3. The third part of the document presents the results of the study. It includes a series of tables and graphs that illustrate the findings of the research. The data shows a clear trend of increasing activity over time, which is consistent with the hypothesis.

4. The fourth part of the document discusses the implications of the findings. It suggests that the results have significant implications for the field of study and may lead to further research in this area.

5. The fifth part of the document concludes the study. It summarizes the key findings and provides a final statement on the importance of the research.

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:59:08 ; Search time 20 Seconds
(without alignments)
1962.509 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLETNPLSPETAMREPAAE.....VOKGRXYELVNIQSLRGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6721	100.0	1334	2	US-08-996-545-2
2	6721	100.0	1334	4	US-09-328-320-2
3	5407.5	80.5	1349	2	US-08-612-734B-2
4	3992	59.4	1307	1	US-08-395-246C-2
5	2682.5	39.9	1408	1	US-08-612-521-2
6	2478	36.9	1280	6	5206352-4
7	2477	36.9	1280	2	US-08-583-276-19
8	2469.5	36.7	1279	2	US-08-784-649A-2
9	2466	36.7	1280	2	US-08-752-447-2
10	2466	36.7	1280	4	US-09-316-167-2
11	2408.5	35.8	1275	3	US-09-120-513-2
12	2408.5	35.8	1275	4	US-09-450-105-2
13	2208.5	32.9	1308	2	US-08-996-644-2
14	2208.5	32.9	1308	3	US-09-352-552-2
15	1991.5	29.6	1302	1	US-08-232-537-2
16	1940	28.9	1263	4	US-09-351-224E-11
17	1757	11.3	791	1	US-08-394-880B-2
18	1746.5	11.1	1622	4	US-08-972-927-6
19	1739.5	11.0	1261	4	US-09-605-785-538
20	1739.5	11.0	1261	4	US-09-439-313-538
21	738	11.0	1621	4	US-08-972-927-3
22	711	10.6	1528	1	US-08-463-092B-6
23	711	10.6	1528	2	US-08-462-109A-6
24	711	10.6	1528	2	US-08-460-907B-6
25	711	10.6	1528	3	US-08-463-179A-6
26	711	10.6	1528	3	US-08-461-384B-6
27	701.5	10.4	580	4	US-09-134-001C-5611

28	701	10.4	327	1	US-08-463-092B-9	Sequence 9, Appl1
29	701	10.4	327	2	US-08-460-907B-9	Sequence 9, Appl1
30	677.5	10.1	1531	1	US-08-463-092B-4	Sequence 4, Appl1
31	677.5	10.1	1531	2	US-08-462-109A-4	Sequence 4, Appl1
32	677.5	10.1	1531	2	US-08-460-907B-4	Sequence 4, Appl1
33	677.5	10.1	1531	3	US-08-463-179A-4	Sequence 4, Appl1
34	677.5	10.1	1531	3	US-08-461-384B-4	Sequence 4, Appl1
35	676	10.1	748	4	US-09-061-764A-19	Sequence 19, Appl1
36	674.5	10.0	1437	3	US-09-061-400-2	Sequence 2, Appl1
37	674.5	10.0	1453	2	US-09-001-273-2	Sequence 2, Appl1
38	674.5	10.0	1453	4	US-08-843-459A-2	Sequence 2, Appl1
39	670	10.0	1228	4	US-09-605-785-537	Sequence 537, App
40	670	10.0	1228	4	US-09-439-313-537	Sequence 537, App
41	669.5	10.0	1531	1	US-08-141-893-2	Sequence 2, Appl1
42	669.5	10.0	1531	1	US-08-463-092B-2	Sequence 2, Appl1
43	669.5	10.0	1531	2	US-08-462-109A-2	Sequence 2, Appl1
44	669.5	10.0	1531	2	US-08-460-907B-2	Sequence 2, Appl1
45	669.5	10.0	1531	3	US-08-463-179A-2	Sequence 2, Appl1

ALIGNMENTS

```
RESULT 1
US-08-996-545-2
; Sequence 2, Application US/08996545
; Patent No. 5928898
GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrd of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-996-545-2

Query Match 100.0%; Score 6721; DB 2; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSPLETNPLSPETAMREPAAETSTEEQASTPHADDEKIIISDLSAPSTATPDKRHP 60
Db 1 MSPLETNPLSPETAMREPAAETSTEEQASTPHADDEKIIISDLSAPSTATPDKRHP 60
```

```

OY 61 KSSSSNNVAVNEVDALLHLPEDEROVLKTOLEEKVNIISFEGMLRYATKMDILIMVTS 120
DB 61 KSSSSNNVAVNEVDALLHLPEDEROVLKTOLEEKVNIISFEGMLRYATKMDILIMVTS 120
OY 121 TICAATAAFTQIRIMLYQSYDEFTYDELTKNLYFYLLGIGEVYVYVSTVGTIYGEHA 180
DB 121 TICAATAAFTQIRIMLYQSYDEFTYDELTKNLYFYLLGIGEVYVYVSTVGTIYGEHA 180
OY 181 TOKIRREYTESILRONIGYFDKIGAGEVTTTRITADTNLIODGISEKVGTLFPLATFYTA 240
DB 181 TOKIRREYTESILRONIGYFDKIGAGEVTTTRITADTNLIODGISEKVGTLFPLATFYTA 240
OY 241 FIITAVKYWKLLALICSSITVALVLTNGGSGQFIITKYSKSLDSYGAGTVAEVVISIRN 300
DB 241 FIITAVKYWKLLALICSSITVALVLTNGGSGQFIITKYSKSLDSYGAGTVAEVVISIRN 300
OY 301 ATAFGQODLAROYEYHDEAEKWKTKNOIYMGFMIGAMFGMLSNYVGLGFMWGSRLVD 360
DB 301 ATAFGQODLAROYEYHDEAEKWKTKNOIYMGFMIGAMFGMLSNYVGLGFMWGSRLVD 360
OY 361 GAVDVODILITVLMALIGSFSILGNVSPNAQAFNAVAALAKITGCTIDROSPIDPYNECK 420
DB 361 GAVDVODILITVLMALIGSFSILGNVSPNAQAFNAVAALAKITGCTIDROSPIDPYNECK 420
OY 421 TLDHEGHIELRNKHIYSPREVTWEDVSLSPMAKKTALVGPSSGSKSTVGLVERF 480
DB 421 TLDHEGHIELRNKHIYSPREVTWEDVSLSPMAKKTALVGPSSGSKSTVGLVERF 480
OY 481 YMPVGTIVLDGHDIKDLKMLRQOISLVSOEPLVFGTITKNHIGLITKYEESD 540
DB 481 YMPVGTIVLDGHDIKDLKMLRQOISLVSOEPLVFGTITKNHIGLITKYEESD 540
OY 541 KYRELLENAKMANAHDTTALPEGETVNGRGFLISGOKORIAIARAVSDPKILL 600
DB 541 KYRELLENAKMANAHDTTALPEGETVNGRGFLISGOKORIAIARAVSDPKILL 600
OY 601 DATSALDTPKSGVVOALEPRAEGRTTVIAHRSTITANITVVLNGKIAEOGTDE 660
DB 601 DATSALDTPKSGVVOALEPRAEGRTTVIAHRSTITANITVVLNGKIAEOGTDE 660
OY 661 LYDRGAYAKLYEAOIRINOKEDALIEDADIEDLTNADIAKITASASSDDGKRTTD 720
DB 661 LYDRGAYAKLYEAOIRINOKEDALIEDADIEDLTNADIAKITASASSDDGKRTTD 720
OY 721 RGTGHSVSSAILSKRPETTPKYSLMTLLKFAVASFNPBEIPYMLGLVFSVLGAGGQPT 780
DB 721 RGTGHSVSSAILSKRPETTPKYSLMTLLKFAVASFNPBEIPYMLGLVFSVLGAGGQPT 780
OY 781 QAVLYAKAISTSLPESQSKLRHADFWSLMFVVGIIQFTTOSTNGAFAVCSRLIR 840
DB 781 QAVLYAKAISTSLPESQSKLRHADFWSLMFVVGIIQFTTOSTNGAFAVCSRLIR 840
OY 841 RARSTAFITLRODIAFFKEENSTGALTSPFSTETKHLGVSQVLTGTLMTSTLGA 900
DB 841 RARSTAFITLRODIAFFKEENSTGALTSPFSTETKHLGVSQVLTGTLMTSTLGA 900
OY 901 IITALAIGKMLALVCISVVPVLLAGCFYRFYMLAQOSKSLAEGSNAFACATSIRT 960
DB 901 IITALAIGKMLALVCISVVPVLLAGCFYRFYMLAQOSKSLAEGSNAFACATSIRT 960
OY 961 VASLTPREVDWEIYHQAQDQRTSLISYLRSSLYASSQALVFVVALGFWTGGTLLH 1020
DB 961 VASLTPREVDWEIYHQAQDQRTSLISYLRSSLYASSQALVFVVALGFWTGGTLLH 1020
OY 1021 HEYDIFRFVCESEILFGASGATVFSFAPDMGKAKNNAEFPRLDRKPOIDNMESEGE 1080
DB 1021 HEYDIFRFVCESEILFGASGATVFSFAPDMGKAKNNAEFPRLDRKPOIDNMESEGE 1080
OY 1081 KLETVGELEFRNVRHRYTPRPOVLRGLDLTVKPGQVVALVGPSCGCKSTTIALLERF 1140
DB 1081 KLETVGELEFRNVRHRYTPRPOVLRGLDLTVKPGQVVALVGPSCGCKSTTIALLERF 1140

```

```

OY 1141 YDAIAGSILVDGKDISKLNINSRSLVSOEPTLYQGTINENILLAGIVEDVPEEFLI 1200
DB 1141 YDAIAGSILVDGKDISKLNINSRSLVSOEPTLYQGTINENILLAGIVEDVPEEFLI 1200
OY 1201 KACKDANIYDFIMSLPEGNTVWGSKGMLSGGOKORVAAIARALLDPKILLDEXTSL 1260
DB 1201 KACKDANIYDFIMSLPEGNTVWGSKGMLSGGOKORVAAIARALLDPKILLDEXTSL 1260
OY 1261 DSESEKVVQAAADPAARGTTTIAHRLSTIQKADVIYVFDQKIVESGTHSELVOKGR 1320
DB 1261 DSESEKVVQAAADPAARGTTTIAHRLSTIQKADVIYVFDQKIVESGTHSELVOKGR 1320
OY 1321 YVELVNLQSLGKG 1334
DB 1321 YVELVNLQSLGKG 1334

```

```

RESULT 2
US-09-328-320-2
; Sequence 2, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrod of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U. S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/328,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-328-320-2
Query Match 100.0%; Score 6721; DB 4; Length 1334;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSPLEFNPLSPETAMREPPESTTEQASTPHADEKTLISDLSAPSSSTTAPADKEHRP 60
DB 1 MSPLEFNPLSPETAMREPPESTTEQASTPHADEKTLISDLSAPSSSTTAPADKEHRP 60
OY 61 KSSSSNNVAVNEVDALLHLPEDEROVLKTOLEEKVNIISFEGMLRYATKMDILIMVTS 120
DB 61 KSSSSNNVAVNEVDALLHLPEDEROVLKTOLEEKVNIISFEGMLRYATKMDILIMVTS 120

```

```

Db 61 KSSNNNAVNEVDALLAHPEDEKQVLTQLEIKVNI SFGMLRVTAKMDILIMVIS 120
Qy 121 TICATAAATFORIMLYQISYDEYDELTKNLYEYLLGIGFEVYVTSVGEITYGEHA 180
Db 121 TICATAAATFORIMLYQISYDEYDELTKNLYEYLLGIGFEVYVTSVGEITYGEHA 180
Qy 181 TOKIREYLESILRONIGYFDKIGAGEVTTTITADTNLIQDGISEKVGTLTALATFVA 240
Db 181 TOKIREYLESILRONIGYFDKIGAGEVTTTITADTNLIQDGISEKVGTLTALATFVA 240
Qy 241 FIATAVYKWKLLICSSITVALVLMGGSSOFTIKYSKSLDSYGAGGVAAEVISSIRN 300
Db 241 FIATAVYKWKLLICSSITVALVLMGGSSOFTIKYSKSLDSYGAGGVAAEVISSIRN 300
Qy 301 ATAFGTODKLAYOEYHDEAEKMGTKNOIYMGFMIGAMGLMYSGYGFMMWSRFLVD 360
Db 301 ATAFGTODKLAYOEYHDEAEKMGTKNOIYMGFMIGAMGLMYSGYGFMMWSRFLVD 360
Qy 361 GAVDVGDIITVIMAILIGSFSLGNSVPAQAFTNAVAAAATJFGTIDROSPLDPYSNEGR 420
Db 361 GAVDVGDIITVIMAILIGSFSLGNSVPAQAFTNAVAAAATJFGTIDROSPLDPYSNEGR 420
Qy 421 TLDHBEGLHLEKNVHTIYSPREYVIMEDVSLMPAGKTTALVGPBGSKSTVGLVERF 480
Db 421 TLDHBEGLHLEKNVHTIYSPREYVIMEDVSLMPAGKTTALVGPBGSKSTVGLVERF 480
Qy 481 YAPVGTVLLDGHDKDLMLRMLROQISLYSOEPVLEFGTIIYKNHGLIGTYEENESD 540
Db 481 YAPVGTVLLDGHDKDLMLRMLROQISLYSOEPVLEFGTIIYKNHGLIGTYEENESD 540
Qy 541 KYRELIENAKAMAHDFITLALPEGETVNGOGFLLSGGOKORIAIARAVSDPKILL 600
Db 541 KYRELIENAKAMAHDFITLALPEGETVNGOGFLLSGGOKORIAIARAVSDPKILL 600
Qy 601 DATSALDTPKSGVOALERAAGERTTIVIAHRLSTITAHNIYVLMGKIAEOGTHDE 660
Db 601 DATSALDTPKSGVOALERAAGERTTIVIAHRLSTITAHNIYVLMGKIAEOGTHDE 660
Qy 661 LVDRGAGYKLYEAOIRINQKEADLEDAEDLTNADIAKITASASSASDLGKPTTD 720
Db 661 LVDRGAGYKLYEAOIRINQKEADLEDAEDLTNADIAKITASASSASDLGKPTTD 720
Qy 721 RGTGHSVSSAILSKRPETTPYSLMTLLKFAVSPNREIPYMLGLVFSVLAGGQPT 780
Db 721 RGTGHSVSSAILSKRPETTPYSLMTLLKFAVSPNREIPYMLGLVFSVLAGGQPT 780
Qy 781 QAVLAKAISTSLPESQYSKLRHADDFWLSMEFVVGIIQFTIOSTNGAFAVCSERLIR 840
Db 781 QAVLAKAISTSLPESQYSKLRHADDFWLSMEFVVGIIQFTIOSTNGAFAVCSERLIR 840
Qy 841 RARSTAFRTILRODIAFPFKENSTGALTSFLETKHLGSGVGLTITIMSTLGA 900
Db 841 RARSTAFRTILRODIAFPFKENSTGALTSFLETKHLGSGVGLTITIMSTLGA 900
Qy 901 IITALAIGKLLALVCISVVPVLLACGFYRPMYLAQFOSRSKLAEGSANFACATSIRT 960
Db 901 IITALAIGKLLALVCISVVPVLLACGFYRPMYLAQFOSRSKLAEGSANFACATSIRT 960
Qy 961 VASLIRERWYELIYAOQDQRTSLISYLRSSLIYASSOALVFCVAGLGTGGLLGH 1020
Db 961 VASLIRERWYELIYAOQDQRTSLISYLRSSLIYASSOALVFCVAGLGTGGLLGH 1020
Qy 1021 HEYDIFRFVCSSEILFGAOSAGTVSFPADMGKAKNAAEFRRLDRKQIDNMSSEGE 1080
Db 1021 HEYDIFRFVCSSEILFGAOSAGTVSFPADMGKAKNAAEFRRLDRKQIDNMSSEGE 1080
Qy 1081 KLETVGEIEFRNVHRYPTRPEQPYLRGLDLVKGQYVALVGPBGCKSTTIALLERF 1140
Db 1081 KLETVGEIEFRNVHRYPTRPEQPYLRGLDLVKGQYVALVGPBGCKSTTIALLERF 1140
Qy 1141 YDAIAGSILVDGDISKLNINRSFLSLVSOEPTLYOGTIKENIILGLIVEDVPEEFLI 1200
Db 1141 YDAIAGSILVDGDISKLNINRSFLSLVSOEPTLYOGTIKENIILGLIVEDVPEEFLI 1200

```

```

Qy 1201 KACKDANIYDFINSLPEGFVTVSGKGMJSGGOKORVATARALRDPKILLDEATSA 1260
Db 1201 KACKDANIYDFINSLPEGFVTVSGKGMJSGGOKORVATARALRDPKILLDEATSA 1260
Qy 1261 DSESEKVOQALDAARGRTTIAVAHRLSTIQADVIYVFDQGIYESGTHSELVOKGR 1320
Db 1261 DSESEKVOQALDAARGRTTIAVAHRLSTIQADVIYVFDQGIYESGTHSELVOKGR 1320
Qy 1321 YVELVNLQSLGKH 1334
Db 1321 YVELVNLQSLGKH 1334

RESULT 3
US-08-612-734B-2
; Sequence 2, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DC1501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1349 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-612-734B-2

Query Match 80.5%; Score 5407.5; DB 2; Length 1349;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;
Qy 5 ETNPLSPETMRPAETSTTEBOASTPHADEKKILSDLSAPSTVATPPADKREPRSS 64
Db 5 ETGASSSEKL-BDLYATLEKGRSTSSGADNEKPRDHSLDITNAPD-----GKKRD 59
Qy 65 SNNAVSYNEVDALIAHLPEDEROVLKTQLEIKVNI SFGMLRVTAKMDILIMVISTICA 124
Db 60 HGRAVDLND-DSLFAHLOEHEKEVKLRQLDAPVSVSFFTLRYASRKDILLIIVSAICA 118
Qy 125 IAA-----ASTFORIMLYQISYDEYDELTKNLYEYLLGIGFEVYVTSVGEITYGE 170
Db 119 IAAGAALPLFTIILGSLASFOGISGTMPYHEFHKLITKNVLYFYLLGIAEEVTVYVST 178
Qy 171 VGEITYGEHATQKIREYLESILRONIGYFDKIGAGEVTTTITADTNLIQDGISEKVG 230

```



```

DB 305 NGIGLFWMSRFLVEGISNFKADVLTIMAAIILGSSYNLGNVAPNGOALSDAANAASKLY 364
QY 404 GTIDROSPIDPVSNEKTLDFHEGHI ELRNVKHIYPSRPEVTWEDVSLMPGKTTALV 463
DB 365 GTIDROSPIDLSDDQGTLEFVGRNVIYLOINIRVYPSRPEVTYAHNDISCTYIPGKTTAFV 424
QY 464 GPSGSGKSTVGLVEFVYFVGRVTLDDGHIDRLRMLROQISLVSOBPVFGTTIK 523
DB 425 GPSGSGKSTIISLEFVYDPVACTIMLDGHDITLNRMLROQMSLVSOBPVFAITIAE 484
QY 524 NTRHGLIGTKYENESDQVRELIEAKAANAHDFITALEPEYETVNGOQGLISGQOK 583
DB 485 NTRGIGISRFKESTYEIKRVEAARAHANADFTALPNGYDTNI--ESFISGQOK 542
QY 584 RIAIARAVSDPKIILDEATSLDPSSEGVQALERAEGRTTYIARLSTIKAHN 643
DB 543 RIAIARAILIKDPIILDEATSLDPSSEKLVQALDKSKGRTTYIARLSTIKAHN 602
QY 644 IYLVNGKIAEQTHDELVDYRGAYRKLVEAQRINEQKADALDADADIEDLTN----- 696
DB 603 IYLVNGKIAEQTHDELVDYRGAYRKLVEAQRINEQKADALDADADIEDLTN----- 696
QY 697 -----ADIATKTASASSDIDGKPTTIDRTGTH-----KVSASALSK-R 736
DB 661 MPEFPKDYPGDESDIYSI--LSDNASDI-----GLHTGEKORVSRMSLSHM 708
QY 737 PREPTKYSIWLTKLPASINRPEIPIYMLGLVPSVLAGGGOPOAVLVKAIISTSLPE 796
DB 709 QAPKEENASVWTLKFKPLASINRPEIPIYMLGLVPSVLAGGGOPOAVLVKAIISTSLPE 768
QY 797 SOYSKLHDADFWLMEFVVGIIQFIQSTNGAFAVCSRLIRARSTAFRTILRODIA 856
DB 769 LEYKRLHDANFQIMIGIYIVYISVQGLTFAVSSKMYRANSQAFVILHODIS 828
QY 857 FPDKENSTGALTSTSTETKHLGVSQVTLGTLMTSTLGAIIITAIAGWKALVCI 916
DB 829 FPDQOEWTLTATLTATLISAGKELTIGISVTLGTLIYVNLVSLAGVLAIVGKALVCI 888
QY 917 SVYVILLAGEFYRPMYMAOQSRKSLAVESANPACATSIPIRVYALTEBDROWELIYHA 976
DB 889 SAVALMCGFVYRPMYMAOQSRKSLAVESANPACATSIPIRVYALTEBDROWELIYHA 948
QY 977 QLDAGOSTLSIYLRSSILYASSQALVFCVAGLFWYGTLLGHHEYDIRFEVCFSEIL 1036
DB 949 QLRQOLKSDILPIYKSSILYASSQALVFCVAGLFWYGTLLGHHEYDIRFEVCFSEIL 1008
QY 1037 FGQOSACTVTSFAPDMGKAKANAEEFRRLDPRKQIDNMSBEKLETVGEIEFRVNH 1096
DB 1009 FGQOAGCTVTSFAPDMGKAKANAEEFRRLDPRKQIDNMSBEKLETVGEIEFRVNH 1067
QY 1097 RYPTPRPOVPLRGDILVVRGQOYVALVPGSGCKSTTITALLERDYDAIAGIILVDGDIS 1156
DB 1068 RYPSRLEQPLRLHNLTIKIGQAFVALVAGSGSKSTTITALLERDYDPLKGVYVDGKNIT 1127
QY 1157 KLININSYRSLSVSOBPVTLIYOGTIKENILGLIYEDVPEEFILKACDANIYDFIMSLP 1216
DB 1128 TLEKSSYRSHALISQBPVTLIYOGTIKENILGLIYEDVPEEFILKACDANIYDFIMSLP 1187
QY 1217 EGFRTVGSAGGMLSGQOKQVARIARALRDLPKIILDEATSLDSESEKVVQALDAAA 1276
DB 1188 QGFRTVGSAGGMLSGQOKQVARIARALRDLPKIILDEATSLDSESEKVVQALDAAA 1247
QY 1277 RGRRTIYVAHRLSTIYOKADVIYVFDQKIVESGHSSELVYOKKGRVYELVYNLS 1329
DB 1248 RGRRTIYVAHRLSTIYOKADVIYVFDQKIVESGHSSELVYOKKGRVYELVYNLS 1300

```

RESULT 5
 US-08-612-521-2
 ; Sequence 2, Application US/08612521
 ; Patent No. 5786463
 ; GENERAL INFORMATION:

```

? APPLICANT: Peery, Robert B
? APPLICANT: Skatrud, Paul L
? APPLICANT: Thornwell, Susan J
? TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
? TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
? NUMBER OF SEQUENCES: 3
? CORRESPONDENCE ADDRESS:
? ADDRESS: Patent Division/AEH
? STREET: Lilly Corporate Center
? City: Indianapolis
? STATE: Indiana
? COUNTRY: USA
? ZIP: 46285
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patent Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/612,521
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Hamilton, Amy E
? REGISTRATION NUMBER: 33,894
? REFERENCE/DOCKET NUMBER: X-9693
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 317-276-3169
? TELEFAX: 317-276-1294
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1408 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-612-521-2
?
? Query Match 39.9%; Score 2682.5; DB 1; Length 1408;
? Best Local Similarity 43.9%; Pred. No. 2,6e-234;
? Matches 597; Conservative 233; Mismatches 449; Indels 81; Gaps 16;

```

[illegible]

```

Db 904 FRTVSVLTQPKFHEHMAOSLQVYRNLSLRKAHIFGIFTFSFTQAMTFYAGCFRFGAYL 963
QY 1018 LGHHEVDLFFRFVCFSEILFGAQSAGTVFSFADPMGKAKNAAEFRRLPRKPIDMWSE 1077
Db 964 VAHKLMSFEDVLVFSNAVVGAMAVGVSSFPADYAKAKISAHIIMIIEKTPILDISYST 1023
QY 1078 EGEKLEVEGEIEFRNHFYPTRPBPQVLRGLDLYVKPGQYVALVPGSCGSKSTYALL 1137
Db 1024 EGLMPTLBEONVTFGEVVFYPTRPDIPLVQGLSEVKKKQTLALVSSCGSKSTYVOLL 1083
QY 1138 ERFYDAAGSILVDGKDISKLINSTRSPLSLVSOEPTLYOGTITENILGLYEDVDPEE 1197
Db 1084 ERFYDPLAGVLLDGEIKRLNQLWMLFAHLGIYSOEPIILDSCSAENIAYGDNSSRVSOE 1143
QY 1198 ELKACDAIYDIFMSLPEGFMTYVSGSGGKMSGGOKORAVATRALRDPKILLDEAT 1257
Db 1144 EIVRAAEANIHAFIESLPRKYSTKYGDKCTQLSGGOKRIARALRVPKHILLDEAT 1203
QY 1258 SALDSESEKVVQAALDAARGRTTIAVAHRLSTLOKADVIYVEPOCKIVESGTHSELVOK 1317
Db 1204 SALDTESEKVVQAEADARGRTCIYAHRLSTLOKADVIYVEPOCKIVESGTHSELVOK 1263
QY 1318 KGRYELVNIQS 1329
Db 1264 KGIYFSMVSVOA 1275

RESULT 7
US-08-583-276-19
; Sequence 19, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSFECTED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carella, Byrne, Bain, Giffillan,
; ADDRESSEE: Cecchi & Stewart
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: DM4.V2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,276
; FILING DATE: 05-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/332,444
; FILING DATE: 31-OCT-1994
; APPLICATION NUMBER: 07/887,712
; FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1280 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
US-08-583-276-19
Query Match 36.9%; Score 2477; DB 2; Length 1280;

```

```

Best Local Similarity 41.9%; Pred. No. 1,le-215;
Matches 533; Conservative 230; Mismatches 449; Indels 60; Gaps 9;
QY 94 EIVNINISFGLMRYAKRIMOLIVIVISTICALAASTFORIMLY----- 137
Db 28 KKKPPTVSFVSFMSFYSNMWLDKLYMVGTLAIIHGAGLPLMLLFGEMTIDIFANAGNLED 87
QY 138 -----QISYEEFY----DELTKNVLVYVIGIGFVTVVSTVGFTITGEHAQKI 184
Db 88 LMSNITRSDINDGPFNNLEEDMTTRAYIYSGAGVLAAYIQVSWCLAAARQIHKI 147
QY 185 REYVLESILRONIGYFDKLAGEVYTRITADPNLIQDISERKGLTTLATAPVYAFIIA 244
Db 148 RKQFHAIMROEIMQFVHDVGEINLRLTDVSKINEVIGDKIMQFOSMATFFGTGVIYV 207
QY 245 YVKWTKLALICSSITVALYILMGSGSQFIYKSKSDSYAGAGTVAEVYSIRNMTAF 304
Db 208 FTRQWKLTLLVILALSPYLIGSAVAWAKLISFTDKELIAYAKAAVAEVLAAIRYIAF 267
QY 305 GTQDKLAKQYEVHLDEAEKMGTKNQIYWGFMIGAMFGIMYSGNGIGFMSGRPLVDGAVD 364
Db 268 GGQKKELERYNKNLEAKRIGIKKAITANISIGAAFLIYASYALAFYGTTLVSGEYS 327
QY 365 VGDILYVLMALIGSFSLGNVSPNAQAFNAVAATAKIFGTIDROSPLDPYNSGKTLIDH 424
Db 328 IGQVLTVEFFSVLIGAFSGQASPSIEAFANARGAAYEIFKIIDNKPSIDYSKSGHKPDN 387
QY 425 FECHIELRNKHYTPSREPVYVMDVSLSMRAGKTALVYSGSGSKSVVGLVERFYVY 484
Db 388 IKGNIEFRNHFYSYRSREVRILKGLNLTQVSGQVALVGNSSGCKSTTVOLMORLXDP 447
QY 485 RGVTLDOHDKIDNLRLKROISLVSOEPLFGFTTYKRNIRHGLIGTKYENESDKVRE 544
Db 448 EGMATVSGODIRTNVRLREIIGVSOEPLVFTIENIRY-----RQNTYDE---- 499
QY 545 LIENAKMANAHDFTALPEGETNVGGRFLSGGOKORIAARAAYSDRPIILLDEAT 604
Db 500 -IEKAVKANAYDITMKIPHFDTLVGERGALSGGOKORIAARAALVRNPKIILLDEAT 558
QY 605 SALDTSRGVQAALERAEGRTTIVAHRLSTLTANINIVLVNKGIAEGTHDEVD 664
Db 559 SALDTESEAVVQALDKARKRTTIVAHRLSTLVNNAVYIGFDDGVVEGNDHELMKE 618
QY 665 GCAYRKIVEAQRINDQKADALADADLTNADIATKTASSASSDIDGRPTTIDRTGT 724
Db 619 KGIYFKLVTGTAGNEVE---LENA-----ADESKEDIALEKSSSDSSILRKST 668
QY 725 HKVSSAALSKRP-----PETPKYSLWTLKLFVASFNPPEIPYMLIGLVFSYLAGG 777
Db 669 RRSYRGSAQDRKLTSTKEALDESIPVPSFWIRIMK---LNTLEMPYFVGVCAIINGCL 724
QY 778 OPTQAVILAKAISTLSLPSQYSKLRHADFMSLMEFFVGIIOFTOSTNGNAAFVCSER 837
Db 725 QPAPALIFSKLIIGVTRIDDEPK-RQNSNLSLFTLALGISFTFLQGFTEKAGEI 783
QY 838 LIRRASTAFPTIIRODIAPFDKEENSGALTSFTFKHLSGVSQVTLTILMTSTYL 897
Db 784 LTKLRKRVYFERSMLRQDVSWDDPKNTGALTTRIANADAQYKKAIGSRILAVITQNIANTL 843
QY 898 GAAIILALAIKWLALVCSIVPVLLACGFYRFYMLAOFQSRKSLAYBGSAFACAEATSS 957
Db 844 GTGIIISFYQMOLITLALAIPIIAGVYEMKLSGQALKDKKELGAGKAIATEALEIN 903
QY 958 IRTYASTLRDQWETIYNAQDAGQRTSLIVRSLLIYASQALVYFCVALGFWYGGTL 1017
Db 904 FRTVSVLTQPKFHEHMAOSLQVYRNLSLRKAHIFGIFTFSFTQAMTFYAGCFRFGAYL 963
QY 1018 LGHHEVDLFFRFVCFSEILFGAQSAGTVFSFADPMGKAKNAAEFRRLPRKPIDMWSE 1077
Db 964 VAHKLMSFEDVLVFSNAVVGAMAVGVSSFPADYAKAKISAHIIMIIEKTPILDISYST 1023
QY 1078 EGEKLEVEGEIEFRNHFYPTRPBPQVLRGLDLYVKPGQYVALVPGSCGSKSTYALL 1137

```

Db 1024 EGIAMPTELEGNTFGEVENVPTPRDIPVLOGLSLEVKKGQTLALVSSCGCKSTVVOQL 1083
 Qy 1138 EREYDAIAGSLVLDGDISKLNINSYRSLSVSOEPTLYOGTIKENILLGIEDVPEE 1197
 Db 1084 EREYDADKAGVLLDGKEIKRLNVOMLRAHLGIYSOEPIIFDCSIAMINIAVGDNSRVVSOE 1143
 Qy 1198 ELIKACKDANIYDFIMSLBEGFTVTVSGKGMISGQOKORAVAIARALBDPKILLDEAT 1257
 Db 1144 EIVRAKENINIAFIESTLPRKSTYKGDGTQOLSGQOKRIARALVROPHILLDEAT 1203
 Qy 1258 SALDSESEKVVQALDAAARGRTTIAVAHRLSTIOKADVIYVFDGKIYESGTHSELVOK 1317
 Db 1204 SALDTESEVAVQALDKAKREGRCIYIAHRLSTIIONADLIVVQNRKVKHGHQOULLAQ 1263
 Qy 1318 KGRYELVNLQS 1329
 Db 1264 KGIYFSMVSQA 1275

RESULT 8

US-08-784-649A-2
 ; Sequence 2, Application US/08784649A
 ; Patent No. 5830697

GENERAL INFORMATION:

; APPLICANT: SIKIC, Branimir I
 ; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
 ; TITLE OF INVENTION: CYCLOSPORIN MODULATION
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 2200 Sand Hill Road
 ; CITY: Menlo Park
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94025
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/784,649A
 ; FILING DATE:

CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sherwood, Pamela J
 ; REGISTRATION NUMBER: Reg. No. 5830697 36,677
 ; REFERENCE/DOCKET NUMBER: 06037/007001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-322-5070
 ; TELEFAX: 415-854-0875

INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1279 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-784-649A-2

Query Match 36.7%; Score 2469.5; DB 2; Length 1279;
 Best Local Similarity 42.0%; Pred. No. 5.3e-215;
 Matches 534; Conservative 229; Mismatches 448; Indels 61; Gaps 10;

Qy 94 EELKAVISPEGLMRYATKKDILIMVSTICATAAATFORIMY----- 137
 Db 28 KKKKPVSVFMSFRYSNMIDKLMVVGTLAAIIGHAGLPLMLVGEEMTDIPRANAGNLED 87
 Qy 138 -----QISDEYF-----DELTKNLYLYFYLLGIGEEVTVVSTGFIYGEHATOKI 184
 Db 88 LMSNITNRSDINDTGFPMLEEDMTIRYAYYSGIGAGVVAAYLIQVFCFLAAGROIHKI 147

Qy 185 REYELSLIKONIGYEDKLGAGEVTTTRITADNMLIODGISEKVGTLTALAEFTVAFITA 244
 Db 148 RKOFPAHAKMOEIGWPDVHVGELNTRLDVDSKINEVIDKIGMFQSGAFTEFGTIG 207
 Qy 245 YVYKRWALICSSITYALVLTMGSSQFIITKSKLSDYGAGGYVAEYISSIRNATF 304
 Db 208 FYNGMKTLTYVTLAISPVLGISAAVMAKILSSFQDKELLAFAKGAVAEVLARIVIAF 267
 Qy 305 GTQDKLAKOYEVALDEAEKGTGKNOIVMGFMIGAMGLMYSNGLFGMWSRFLVGAAD 364
 Db 268 GGGKTELEERNKMLEAKRIGIKAITANISIGAAFLITIASALAFVGTITVLSGEYS 327
 Qy 365 VGDILTVLMAILLIGSFLGNSVSPNAOAFYNAVAAAKIFGTIDRQSPLDIPYSNEGTLDH 424
 Db 328 IGGVLTIV-FSVLLIGAFSVGASPSIEAFANRGAAYEIFFIIDNKRISIDYSKSGHKPDN 386
 Qy 425 FEGHIELRNKHIYPSRPETVMEVDSLSMPAKTALVQSSGSGSTVGLERYMYV 484
 Db 387 IKGNLEFRNVHESYPSRKEVKILGLNLVQSGQVAVLGVNSCGKSTVQOLMQLYDPT 446
 Qy 485 RGTVLLDGHDKDLNLRMLRQOISLVSOEPEVLEGTITIKRIRHGLTGYENESDEKVE 544
 Db 447 EGMVSVQDQDIRINVRFLREILGIVSOEPEVLEFATITAIENIRG-----RENTYME---- 498
 Qy 545 LIEAANKANANDFTALPBEGYTNVGOEFLISGQOKORIAIARAVSDPKILLDEAT 604
 Db 499 -IEKAVKENAVDYFIKLPKFPDLVGERGAOLSGQOKORIAIARALVNPDKILLDEAT 557
 Qy 605 SALDTESEVAVQALDAAARGRTTIAVAHRLSTIOKADVIYVFDGKIYESGTHSELVOK 664
 Db 558 SALDTESEVAVQALDAAARGRTTIAVAHRLSTIOKADVIYVFDGKIYESGTHSELVOK 617
 Qy 665 GGAYRKLVEAQRINEQKADDAEDLJTNADIKIRITASSASSDDLDPKPTIDRTGT 724
 Db 618 KGIYFLVLTMOFAGNEVE--LENA-----ADESKSEIDALEMSNDSRSSLLIKRST 667
 Qy 725 HKVSSAILSKRP-----PETTPKYSIMTLILKFAVASTNRPEIPMLIGLVSVIAGCG 777
 Db 668 RRSVRGSOAQODRLSTYKALDESIPVPSFWRMK---INLEMPYFVVGVFCALINGSL 723
 Qy 778 OPTQAVLVYKAKISTLSLPESQYSKLRHDDFMELFVVGIIOTIQTSGNAAFAVNCSSR 837
 Db 724 QPFAIIFRSKIIGVFTRIDDPETK-RQNSNLSLELALGIIISFTIIFLOGFTFGAGEL 782
 Qy 838 LIRARSTAFRTLLRODIAFFDEKENSATGALSTFETKHLSGVSGVLTGTLMTSTL 897
 Db 783 LTRKLRVYFRSMLRQDVSFMDPKNTGALTRLRLANDAAQVGAIGSLAVITQIANL 842
 Qy 898 GAATITIALAIGKIALVCSIVPVLLACGFYRPMIAQFOSRKLAYESGANFACATSS 957
 Db 843 GTGIIISFTYGMQTLTLLLAIVPIIAAGVEMKMSGQALDKKLEBAGKATATEAIE 902
 Qy 958 IRTVASLTEREDVWEIYHQAOLDAGRTSLISVRSLSLVAASSQALVFCVAGFWGTL 1017
 Db 903 FRTVVSLTQDQKEHMYTAOSLOVPYRNSLRKAHIFGITTSFTQAMMYFSYAGCFRGAVL 962
 Qy 1018 LGHHEYDIFREFYCFSEILFQASQAGTFSFAPDMKANAFAEFLRDRKPOIDNMSE 1077
 Db 963 VAHKLMSPEDVLLVFSAVVAGMAVGVSSFAVDYAKAKISAHIMIIETKPLIDSYST 1022
 Qy 1078 EGKELTVEGEIEFRVNRHRYPTRPQPVLRGIDLTVKGOYVALVGPSCGCKSTIALL 1137
 Db 1023 EGIAMPTELEGNTFGEVENVPTPRDIPVLOGLSLEVKKGQTLALVSSCGCKSTVVOQL 1082
 Qy 1138 EREYDAIAGSLVLDGDISKLNINSYRSLSVSOEPTLYOGTIKENILLGIEDVPEE 1197
 Db 1083 EREYDADKAGVLLDGKEIKRLNVOMLRAHLGIYSOEPIIFDCSIAMINIAVGDNSRVVSOE 1142
 Qy 1198 ELIKACKDANIYDFIMSLBEGFTVTVSGKGMISGQOKORAVAIARALBDPKILLDEAT 1257
 Db 1143 EIVRAKENINIAFIESTLPRKSTYKGDGTQOLSGQOKRIARALVROPHILLDEAT 1202
 Qy 1258 SALDSESEKVVQALDAAARGRTTIAVAHRLSTIOKADVIYVFDGKIYESGTHSELVOK 1317


```

GENERAL INFORMATION:
APPLICANT: Mechelner, Eugene
APPLICANT: Roninson, Igor B
TITLE OF INVENTION: Methods and Reagents for Preparing and
TITLE OF INVENTION: Using Immunological Agents Specific for P-glycoprotein
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff Ltd.
STREET: 300 South Wacker Drive, Seventh Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/316,167
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/752,447
FILING DATE: 13-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: No. 6365357nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 95,1121
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-9808
SEQUENCE CHARACTERISTICS:
SEQUENCE FOR SEQ ID NO: 2:
LENGTH: 1280 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-316-167-2

Query Match 36.7%; Score 2466; DB 4; Length 1280;
Best Local Similarity 41.7%; Pred. No. 1,1e-214;
Matches 529; Conservative 233; Mismatches 454; Indels 52; Gaps 9;

```

```

DB 448 EGVAVSDGODIRINVRFLLEIGVVSQEPVLATTAENIRG-----RENTMDE--- 499
QY 545 LINAAMNAHDEFTALPEGYETNANGORFELVSGGOKORAIARAVSOPKILLDEAT 604
DB 500 -LEKAVKENAYDFIMKLPKFDPLVGERGAQSLGGOKORAIARAVSOPKILLDEAT 558
QY 605 SALDTKSEGVQAALERAAGRTTIVIAHRLSTIKTAHNVIVLVNKGIAEOGTHDELVDYR 664
DB 559 SALDTSEAVVQVALDKARKGRITIVIAHRAFAVRNADVAGDDGVYKGNHDELMKE 618
QY 665 GGAIRKLVQAEQRIHQREADADADADADADADADADADADADADADADADADADAD 724
DB 619 KGIFYKLVYVQTAGNEVE---LENAADESKSEIDALEMSNDSSIRRSRRSVRGS 675
QY 725 ---HKSVSALISKRPETPKYSLMTLKFVASFNRPELIPWLVGIVSVLAGGGOPNQ 781
DB 676 QAOHRKLT---KEALDESLPPVSEFRIMK---LNTPEPPIVGVFCALINGGLQDPNF 728
QY 782 AVLYAKAISTLSLPESQSKLRHADFWSLMFFVVGITPITOSTGMAFAVCSERLIR 841
DB 729 AITFSKIGVETRIDDERK-RQNSNLSFLFLALGISITPFLGFTFGKAGELLTR 787
QY 842 ARSTAERTILRODIAFPDKENSTGALTSFSTETKHLSSVSVTIGTITMSTTGAAI 901
DB 788 LRYMVRSMRLQDVSFHPDKNTGALTURLADAQVKAISRLAVIYQNTANLGTGI 847
QY 902 IIALAIGMKLALVCISVAPVLACGFYRFMYLQFOSRSKLAVESANFACENTSSIRV 961
DB 848 IISFYGWQTLTLLLAIVPILAIAGVEMKFMFGALKKKELEGAKITAEIENFRY 907
QY 962 ASLTRERDWEIYHAQDAQGRSLISLVSSLLVASSQALVFVYALGFWYGTLLGHH 1021
DB 908 VSLTQEQKFEMHYAQSLQVYRNSLRKAHIFGITFPTQAMMFYSAQCFRGAIVAHK 967
QY 1022 EYDIFRFVQFSEILLAGOAGGVFSPADPMGAKNAAEFRLEPDRKQIDMWSDEGR 1081
DB 968 LMSFEDVLVFSVAVFGAMAVGVSSFPADYAKAKISAHIMIIEKTPIDISSTEGLM 1027
QY 1082 LETVEGEIEFRNHFRRPPEQPVLRGLDLPKPGQYVALVPSGSGKSTTALLERFY 1141
DB 1028 PNLBEGVNTGEVYFNTPRPDIPVQLGSLERKKGOTLAVSSGCKSTVQLLERFY 1087
QY 1142 DALAGSLVDGKDISKLINYSRFLSVSQEPFLYOGTITENILGLIYEDVPEFLIK 1201
DB 1088 DPLAGKVLIDGKEIKRLNQMWRALHIGIVSQEPILFDCSIAENIYAGDNGRVSQEBYR 1147
QY 1202 ACDANAYDIFMSLPREFNTVSGKMSGGOKORAIARALRDKILLDEATSALD 1261
DB 1148 AAKEANIHAFIESLPNKYSTKVGDKTQLSGGOKORAIARALRDKILLDEATSALD 1207
QY 1262 SESEKVVQAALDAARGRTTIVIAHRLSTIKQADVIVFPOGKIVSGTSELVOKGRY 1321
DB 1208 TESEKVVQEAALDAARGRTTIVIAHRLSTIKQADVIVFPOGKIVSGTSELVOKGRY 1267
QY 1322 YELVNIQS 1329
DB 1268 FMSVSVQA 1275

RESULT 11
US-09-120-513-2
Sequence 2, Application US/09120513
Patent No. 6025160
GENERAL INFORMATION:
APPLICANT: Brun, Kimberly
APPLICANT: Chenevry, Richard
APPLICANT: Ellens, Harma
APPLICANT: Field, John
APPLICANT: Yue, Lin
TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
TITLE OF INVENTION: SEQUENCES ENCODING RAT MDRI2B AND
TITLE OF INVENTION: SCREENING METHODS THEREOF

```


Db	140	ILYVYVIGIARVLVSTYVNTLLTYAAYRIVRYNRIRHAYUKAALUSGEVAYVYDGGSSGSI	199
Qy	212	ITADPNIIDOCISERKGLITLATALVTVAFIIAYKYMKLALISGSTIVAYLIMGSSQ	271
Db	200	ATSMGKILQAGASQIKGILFQGLAAVVTLSLRMKCMKMLTLCICIPVATIGTGVVAA	259
Qy	272	FIIRKSKSLDSYGAGSTVAEEVYSSIRNNAFAGQODKLAKQEVHLEDEAKWGKKNQY	331
Db	260	VEAGHETRIQIIHAQANSFEGITLACVAVNHAQMGDSIVRFDEYLEAHAVGKKISPL	319
Qy	332	MGFIMAGFGLMYSNYGLCFWMMGSRFLYDGAV-DVGDITVYLMALILISFSIGSNPAQ	390
Db	320	LGILFSAEVYIIYVIGYGLAWQOIHMEFGGEIGTADIFTVLLSVIANSINTLLAPSI	379
Qy	391	AFTNAVAAAANKFCTIDROSPLDPYSNEGKTLDFHEGHIELKNVKNHYPSPREVTMEDV	450
Db	380	EFSSRAASAAQLEFLIDRESEINPYKEGIEDEPERVJGDVLEENVTFSYPTRGITVDLNF	439
Qy	451	SLSMPAKTALVGPSSGSGSTVGVYVERFVPMYRGSTVLLDGHDKDINTLRMLRQISLY	510
Db	440	SLKVPARKVTLVGOSSGSGSTVGLLEKRYNFTSGAIRDGNLISELVGCLRNRYLV	499
Qy	511	SOEVLGTTYYKNIIRGLITGYEENESBDKYELLEENAKANAHDFETALPEGEYFNV	570
Db	500	QOEPEVLFGSVFENDIRKGLVGTWENASKEQEMERQOEAKLATAYHEFISLTDOTDLI	559
Qy	571	GORGELLSSGQOKRIATARAIVSDPKILLIDERTSALDPTKSGVYQOALERRAEGRTIV	630
Db	560	GERGGLLSGQOKORVATARSVQSPVYLLIDERTSALDHAETIYQAKLDAKEBRTIYV	619
Qy	631	IAHLSTITKTAHNIVYVKNKIAQDGHDELVDGGAAYKKYLEAQR--INBEKEDALDE	688
Db	620	IAHLKATITRKADNIVVASKCHIVEQGHESLAKQDSVYGLVKIQLAIVNASAHNDVBE	679
Qy	689	ADADLEITNADIKIKITASASDSDLDCKPTTIDTGYCHKSVSAILSKRPETTPKYSIMT	748
Db	680	GEGEDVALLBETE-----TAVIRYTP--SIRGRMSIKNDROYENHKHMD	722
Qy	749	LKLFVASFNR--PEIPMYLIGLVFSVLAGGO-PTQAVLYAKAISTLSIP-ESQYSKLRH	804
Db	723	MLAALAYLVRECPREKMAKAYLVYLLGCLIGCGCAMYPCQAILMSRYEVEFTLSDGAMDK---	779
Qy	805	DADWMSLMEFPVYGIQITPTOSTGCAFPVCSSELIIRASTAFRTILRODIAFPPKEENS	864
Db	780	-GDYFASLVLYLAAGCLICYLAVGYANTLQIHLISMPFRLILHDLRODIOFPREBNT	838
Qy	865	TGALTSELTSTETKHLISGVSVCTGLTILMTSTYTGAILIITALAIGKVLAVCI-SVVPYL	923
Db	839	TGALYSRIDSPIHLELMGYNALVYIAVLOYVYCGIILAIASFMKLGLVYVGGITPELV	898
Qy	924	ACGEYREFYMLAQFQSRSLAYEGSANFACBATSIIRTVASLTRERDWEIYHAQDAQR	983
Db	899	GAGWVRIRIVDSRDLRQKSKRYGTSSSIASPAVNAIRHYESSLAIEETVLRYETEDHAYS	958
Qy	964	TSLSIVLRSLSILVYASQALVFCVYALGFYTGGLLGHREYDIFRPVYCGSELLFQAQSG	1043
Db	959	SSVYPMATMTCFGLYQCIETFWQALGFYWGCVLVSIGETSMVSEFVAFVLYFPGAQASA	1018
Qy	1044	TVFSPADPMCKAKNAAEFRRLFDKRPQIDMNSBESEKLETEVEGETEFENYFRPTPE	1103
Db	1019	QLFQWMSYITKGINATYNTIAMLHLODPTVAETPERENHDKRGSAPLAMDNVFSTPLRPD	1078
Qy	1104	QPVLRGIDLTVKPGOYVALVPGSGCKSTTIALLEFRYDALAGSILVDEKDISKININSY	1163
Db	1079	APIILKGVMLKINKGOFIAPFVGGSSCKSTMIAMLEFRYDPTGSIITIDASTIDINPISY	1138
Qy	1164	RSFLSVISQGEPTLYQGITIKENILGIVEDVYPEEFLIKACKDANIYDFMSLSPBEGFNYY	1222
Db	1139	RNIVALVQOEPTLRYQGITRONISLGAVKSVSQEOIESALRAMANWDFVSSLPDGIYTPA	1198
Qy	1224	GSKGMSLSSGGOKORVATARALIRDPKILLIDERTSALDSEKVVQOALDAADAQR--GRYT	1281
Db	1199	GSQGSQSLSSGGOKRIATARALIRDPKILLIDERTSALDTESEKIYQKALEGARGDRIT	1256

[illegible]

```

OY 451 SLSMPAGKTTALVGPSSGKSTVVGIVREYMPVRCSTVLDDGHDIDLMLRQOISLV 510
    11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 440 SLKVPAGKVTALVGGSGSKSTVGLLEMYNPTSGAIRLIDGNLISELVNGMRRVRVLY 499
OY 511 SOEPVLFGTITKINRHHGLIGTKEVNESEDKVLEIENAKMANADFTALDEGETNY 570
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 500 QOEPVLFQSSVDNITKGLVGTWENASREOMERYOEAKKLAYAHEFTSELTDGDTL 559
OY 571 GORGLLSSGOKORAIARAVVSDPKILLDEATSDLTRKSEGVQOALERAEGRTTY 630
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 560 GERGLLSSGOKORVAIARAVVSOQPKILLDEATSDLPDAETIYQKALDKAAEGRTTY 619
OY 631 IAHRLSTITAHNTIYVLVNGKIAEOGTDELDVDRGATKRYEAQR--INEOKEADALD 688
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 620 IAHKLATIKAKANIYVMSGHVIEOQTHESLAKDGVAGLVKIQMLAVNASAHADVNE 679
OY 689 ADAEDLTNDIAKIKTASASASDLDGKPTTIDRTGTHKSVSSAILSKRPETPKYSLWT 748
    11 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 680 GGEDEVALLLEVIE-----TAVTRYPT--SIRGRANSIKDRDYENIKHMD 722
OY 749 LUKVASFNR--PEIRPYMLIGLVSVLAGGO--PQOAVLYAKAISTLSLP--ESQYSKLH 804
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 723 MAAALAYLVRECEPELKMAYLVYLLGGLGCAMYPGQAILMSRYVEVFTLSGAMLDK--- 779
OY 805 DADFSIMFEVVGIIQFTQSTNGAFAVCSERLIRRASTAFRTILRODIAFPDKEENS 864
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 780 -GDFYASMLIVLAAGCLICYLVAGVATNTIAOHLSHMFRLILHMLRODIOFDEHEMT 838
OY 865 TGAITSFELSTFKLSGVSGLTITMTLGAIIITAIIGWLAIVCI--SVYVLL 923
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 TGAIVSRIDSYPHAILLELMGYNIAVYVAVLOYVTCGILALFMSKGLVYVFGGIPVLY 898
OY 924 ACGETRYMIAFOQSRKSLAYEGSANFACATSSIRTVASLTRERDWEIYHAQDQGR 983
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 839 GAGMYRIRDSRLDQTSKRYGTSSISASEAVNAIRTVSSLAIEETVLRRYTEELDHVS 958
OY 984 TSLIVSRSSLLYASQALVEFCVALGFWYGGTLGHHEDYDFRFFVCSSEILFGQSKG 1043
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 939 SSVKPMAMTICFGLQCEIYFWQALGFYGCRLVSLGETSMVSEFVAFLSVFFAQASA 1018
OY 1044 TVFSEAPDGMKANAARLFRDRKPOIDNMSSEGEKLETEGEIEFENHFRYPTRBE 1103
    1019 QLFQWSTSTTKINMTNYIAMLHOLQPYRRETPENHDKPGSGCAPLAMDVAFSTPLRD 1078
OY 1104 QPVRLGLDLYVPGQYVALVGPSSGCKSTTALLERFYDALIAGSILVYDKDISKINISY 1163
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1079 APLIGVNMKIRKGGQFIKGVSSGCKSTMIAMLERFYDPTGSIITIDASTLIDIMPISY 1138
OY 1164 RSFLVLSOEPFLYOGTIRENLLIGVEDVPEEFLIKRCKRANITYDFINSLPEGNITY 1223
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1139 RNTVALVOOEPFLFOGTIRDNISLGDVAKSVDEQIESALRAMAMDFVSSLPOGLIYTPA 1198
OY 1224 GSKGMLSGGOKORVAIARALDRPKILLDEATSDALDESEKVVYQALDAALAR--GRTT 1281
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1199 GSGGQSLSGGQORAIARALIRDRPKILLDEATSDALDESEKVIQKALEGAARODRILT 1258
OY 1282 IAVARLSTICKADVIYVFDQKIVESGTHSELVQKRGRYELVNLQSL 1330
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1259 VAVARLSTICKADVIYVFDQKIVESGTHSELVQKRGRYELVNLQSL 1307

```

```

RESULT 15
US-08-232-537-2
: Sequence 2, Application US/08232537
: Patent No. 5516655
: GENERAL INFORMATION:
: APPLICANT: Peery, Robert B.
: APPLICANT: Skatrud, Paul L.
: TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:

```

```

: ADDRESS: Eli Lilly and Company
: STREET: Lilly Corporate Center
: CITY: Indianapolis
: STATE: Indiana
: COUNTRY: USA
: ZIP: 46285
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/232,537
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Plant, Thomas G.
: REGISTRATION NUMBER: 35,784
: REFERENCE/DOCKET NUMBER: X9212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 317-276-2459
: TELEFAX: 317-276-1917
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1302 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-232-537-2

```

```

Query Match 29.6%; Score 1991.5; DB 1; Length 1302;
Best Local Similarity 35.3%; Pred. No. 1.7e-171;
Matches 471; Conservative 264; Mismatches 521; Indels 77; Gaps 23;

```

```

OY 43 LSPSSTATPADKEHRRPSSSSNNNAVSVN-EYDALIAHLPED-----EROYAKTDL 93
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MSSADGRRTVSSDYGH-PPPSMANEADHNEFTTINSQPSMWHTEFVFSKRIQOQL 59
OY 94 EIKIVNISFPGILMRVAT-KMDILINIVISTICAIAAASFQRI--MLYQI--SYDERYDEL 148
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 GKNPFKSYLDLFKLVDNDAKSAVLAAGLILAIAGCCPLPIIGYFGQITTSFPREDVL 119
OY 149 TKNVLYFVYLGIEFYTVVSVGVFTYGEHATQKIREYLESILRONIGYDKLAGEVY 208
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 RDRLVQLVGACGYFVITVGYAIVAGLTGKISRFRRETLVRLGLGEQAYPD-IKDPDI 178
OY 209 TTRIPADTFLIDGISEKGLTFLTALAFVPAFTIAYVYKWLALICSTIY---ALVIT 265
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 TMLTEKIALIDIGISEKGIQISYIVAAFIYVGLIMAKLTGILFAVPIPLMLIYV 238
OY 266 MGGGQFIIKYSKKSLSIDSYGAGTVAEEVYSSIRNATAFGTODKLAKOYEHDAEKWG 325
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 V--GSSRIAKYTKAATEYEAAGRIAESAIHAVKVVOAGMAENLSKEHRLKLSARA 296
OY 326 TKNQIVMGMTIAMGIMVSNVGLGFWMGSRFLVD-GANDVDILTYMALITGSFSGN 384
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 297 IRRSVSAEMGLVLYFTAYSANALAFWEGSRLAASGSNNACTVAVAVFLLIDASFVVGQ 356
OY 385 VSPNQAETFNANAAAKIIGTIDR--QSPIDPYSNGK--TLDFHFGHLELRVYKHIYSR 441
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 357 RGPFLGSPATIAAAGESEYELIINHPOSELINYSKGEQETESDMKADLVFRNVTYTPAR 416
OY 442 PEVTVMEDVSLMPAGKTTALVGPSSGKSTVVGIVREYMPVRCSTVLDDGHDIDLMLR 501
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 417 TSARALEEMSLILKAGOMNAIYVTSGCCSTLVSLILRLYDISSQOLITGSHDIDPNVR 476
OY 502 WLRQOISLVSOQPVLPFTIYKINRHHGLIGTKEVNESEDKVLEIENAKMANADFTITA 561
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 SLRKTYTALVDOSVLEFSGSVENISYGL---GEHSLSDVYLERCTEAKAANL-DFYDF 532
OY 562 IPEGYETVNGQGLF-LSGGOKORVAIARAVVSDPKILLDEATSDALTRKSGVQOALE 620
    111111 : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 533 LPQGIHTRIGNGYISLSGGQORICTARALVKKPALLLDLDEPTALDANSGLITMDAVK 592

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:58:28 ; Search time 57 Seconds
(without alignments)
2249.883 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLFTNPLSPETAMREPAE.....YOKKGRYELVNLQSLGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3992	59.4	1307	2 T30882	multidrug resistan
2	3068.5	45.7	1362	2 T41534	leptomycin B resis
3	2682.5	39.9	1408	2 T43261	multidrug resistan
4	2493.5	37.1	1287	2 S55692	multidrug resistan
5	2489	37.0	1280	1 DVH01	multidrug resistan
6	2485	37.0	1276	1 DVH1C	multidrug resistan
7	2471.5	36.8	1276	1 A34786	multidrug resistan
8	2453.5	36.5	1276	1 DVMS1	multidrug resistan
9	2392.5	35.6	1277	2 JH0502	multidrug resistan
10	2387	35.5	1276	1 DVMS2	multidrug resistan
11	2379.5	35.4	1279	1 DVH03	multidrug resistan
12	2357	35.1	1278	2 A16466	P-glycoprotein - r
13	2357	35.1	1281	2 T48123	P-glycoprotein iso
14	2335	34.7	1104	1 DVMS1A	multidrug resistan
15	2232.5	33.2	1321	2 T42228	P-glycoprotein sis
16	2230.5	33.2	1321	2 T42842	bile salt transpor
17	2203.5	32.8	1278	2 E86155	probable ABC trans
18	2181.5	32.5	1286	2 T02187	probable ABC trans
19	2173.5	32.3	1275	2 T31073	multidrug resistan
20	2166.5	32.2	1292	2 T48007	P-glycoprotein hom
21	2146	31.9	1230	2 E85023	probable P-glycopr
22	2145	31.9	1289	2 D87789	protein C346.4 [1
23	2135	31.8	1294	2 T19982	hypothetical prote
24	2132	31.7	1229	2 F86155	probable ABC trans
25	2128	31.7	1229	2 D85023	P-glycoprotein-lik
26	2128	31.7	1229	2 T52319	P-glycoprotein-lik
27	2106.5	31.3	1323	2 H85202	hypothetical prote
28	2105	31.3	1302	2 S30327	multidrug resistan
29	2098.5	31.2	1310	2 S30328	multidrug resistan

30	2091.5	31.1	1283	2 A47377	multidrug resistan
31	2079.5	30.9	1245	2 G86404	probable P-glycopr
32	2046.5	30.4	1327	2 T21268	hypothetical prote
33	2035	30.3	1247	2 F86405	probable P-glycopr
34	2024.5	30.1	1286	2 A42150	P-glycoprotein ppg
35	2010.5	29.9	1222	2 T14805	hypothetical prote
36	1998.5	29.7	1233	2 T04251	multidrug resistan
37	1991.5	29.6	1302	2 T06165	multidrug resistan
38	1989.5	29.6	1232	2 T06165	multidrug resistan
39	1988.5	29.6	1321	2 T23476	hypothetical prote
40	1985.5	29.5	1321	2 S27337	multidrug resistan
41	1971	29.3	1302	2 A41249	multidrug resistan
42	1960	29.2	1318	2 T21266	hypothetical prote
43	1946.5	29.0	1316	2 B86240	protein F20B24.12
44	1934.5	28.8	1268	2 T22094	hypothetical prote
45	1933	28.8	1291	2 T21267	hypothetical prote

ALIGNMENTS

RESULT 1

multidrug resistance protein 1 - Aspergillus flavus
C:Species: Aspergillus flavus
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T30882
R:Tobin, M.B.; Peery, R.B.; Skatrud, P.L.
Gene 200, 11-23, 1997
A:Title: Genes encoding multiple drug resistance-like proteins in Aspergillus fumigat
A:Reference number: Z20917; MUID:98036972; PMID:9373135
A:Accession: T30882
A:Status: preliminary; translated from GR/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1307 <TOB>
A:Cross-references: EMBL:U62931; NID:92673946; PID:92673947; PIDN:AAB86655.1
C:Genetics:
A:Gene: mdrl
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homolog

Query Match	Best Local Similarity	Score	DB 2:	Length
Query	60	PKSSSSNNAVSNEVDALIAHLPEDEROYKTOLEIVNVSFEGLMRYATRKMDLIYVI	119	1307;
Db	13	PKSPETGTTT-----GHSVSHAEEVLDROQLTPVSQIGFETRYATRKMDVAILFG	64	
Query	120	STICAIAA-----ASTFORIMLYQISDEYDELTKNVYFYVYIGIGERTV	165	
Db	65	SALAAAGGALPLETVLFGRUTSTFPDLATHRTYDHFHHETLNVVYFYLGAAREVA	124	
Query	166	VYVTSVGFYVGEHATQKIREYVESLIRONIGYFDKLAGETVTRITADTNLIODGISE	225	
Db	125	IYLAIVGFYVIGDHYVQOIRVEYEQAILRONAFEDTLAGETVTRITADTNLIODGISE	184	
Query	226	KVGLITLTAFFVFAFIYAYVYKWLALICSTIYALVLMGGSGQFIIRKSKSDSYG	285	
Db	185	KVGLATGLSTFVFAFIYAYVYKWLALICSTIYALVLMGGSGQFIIRKSKLEVOG	244	
Query	286	AGGTVAAEVVISIRNATAGTODKLAKOYEVLDAEAKMGTKNOIYMGFMIGAMGLWYS	345	
Db	245	RGASMAEDILDSIRIVAFPMQOETLARKYESLAKNABGPKMSKTIYFALMGALLCIWYL	304	
Query	346	NYGLGFMMGSRPLVDGA--VDVGDLIVYALMALIGSFSLGNVSFNAQAFYNAVAAARIF	403	
Db	305	NYGLGFMMGSRPLVDGA--VDVGDLIVYALMALIGSFSLGNVSFNAQAFYNAVAAARIF	364	
Query	404	GTIDQSPLDYVSNQKTLDFEHEGIELRANKHYIPSPREYVYMDVSLSPAGCTTALV	463	
Db	365	GTIDQSPDLALSDQKTLDFEHEGIELRANKHYIPSPREYVYMDVSLSPAGCTTALV	424	
Query	464	GPSGSGKSTVVGIVVERFYMPVAGTVLIDGHDIKDLNRLWLROQISLVQEPVLFTGTIYK	523	

[illegible]

QY 694 LTNADIKIT-----ASSASSDDGKPTTIDRGTGHSVSAILSKRP-----E 739
D 698 DEDMDASLAPMWSHNTDPTDNLNNKNEKDNVFEEDKTLQHVASEIIVPNLPADVGLN 757
QY 740 TTPKYS-----LMTLLKRVASPNRP--EIPYMLGLVPSVLAAGGQPOAVLXA 786
D 758 EEPKSKSKSKNNHNSLTALMTFHSFVKMTIICLLIGLISMGGAAPVOAAVFA 817
QY 787 KAISTLSLPESQSKLNHDAFWSLAFVVGIIQFITOSTGCAFAVCESEIRVARSTA 846
D 818 RFLNIEF--DLSSIDFLHKNVFAVYVWLLALVQFPAVAINFMYATAMEAVLORIKYL 875
QY 847 FRILLRODIAFDKENSTALSTFSTERKHSVSGSVTLTMTSTLGAATIALA 906
D 876 FRLLRQDVFEFFRESEVTAITSTKIQSLGSLGSPFQILITNIIISTILSLA 935
QY 907 IGKMLAVCSVYPVLLACGFYRYMLAQOSRKLAVBESANFACERTSSIRVASLTR 966
D 936 TGMKGLVLTSTSVIITTAQYVRALDQVQEKLSAAKKSMAACESTSAIRVASLNR 995
QY 967 ERDVMETIYHAQDAGRTSLISVRSLSLVAASSQALVFCAVAGFWGTLGHEDIF 1026
D 996 EENVFAFCOSLIRPGRESALASLGSLEFSAAGVTFLLNALTFWIGSTLMKGETNIY 1055
QY 1027 REFVCSSEILFGAQSAGTVESFADMGKANAFAFRRLDRKQIDNWSEGEKLTVE 1086
D 1056 QFYTCFAIYFQIOAGQFPGYSADYTKAKAAGEIKYLSSEKRPIDTWSTEGKVSILQ 1115
QY 1087 -GLEEFNNVFRYPTRREGVYLRGLDITVAPGVVALVYSGCGCKSTTIALLEFYDAIA 1145
D 1116 SAALIEFQVSESYTRKHIVKRLNLTVRPGVAVAGSSGCKSTTIGIEFHYCDN 1175
QY 1146 GSILVNDKDISKLININSYRFLSVSEPTLYOCTIKENILLIGVEDVPEEFLKACKD 1205
D 1176 GAVLVQSVNVRDYNINRYQIALVYSEPTLYOCTVRENTIVLG--ASNDVSEEMIEACK 1234
QY 1206 ANIYDFMSLPEGFNVYVSGKGLSGGQKORAVIALRLDRPKILLDEATSLDSESE 1265
D 1235 ANIHEFTLPGNYNLTGCGKSSLSGQKORAIARALIRNPKILLDEATSLDSESE 1294
QY 1266 KVVQALIDALARGRTTAVAHRLSTIOKADIVYFDDGKIVESGTHSELVQKRYEYL 1325
D 1295 KVVQALIDALARGRTTAVAHRLSTIOKADIVYFDDGKIVESGTHSELVQKRYEYL 1354
QY 1326 NLQSLGK 1332
D 1355 VEOGLNK 1361

RESULT 3
T43261
multidrug resistance protein 1 - fungus (Filobasidium floriforme)
C:Species: Filobasidium neofomans; Cryptococcus neoformans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 11-May-2000
R:Accession: T43261
R:Thornhill, S.J.; Peery, R.B.; Skatrud, P.L.
Gene 201, 21-29, 1997
A:Title: Cloning and characterization of CneMDR1: a Cryptococcus neoformans gene encoding
A:Reference number: 222374; PMID:98072425; PMID:9409767
A:Accession: T43261
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1408 <THO>
A:Cross-references: EMBL:U62929; NID:g2668552; PID:g2668553; PIDN:AAC49889.1
C:Genetics:
A:Gene: mdr1
A:Map position: IV
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; membrane protein; P-loop

Query Match 39.9%; Score 2682.5; DB 2; Length 1408;
Best Local Similarity 43.9%; Pred. No. 3; 5e-154;

Matches 597; Conservative 233; Mismatches 449; Indels 81; Gaps 16;
QY 33 AADKKLIDLSAPSSPTA--TPADKREHPKSS--SSNNAVSVNEVDALIAHLPEDROYL 89
D 69 AAEGRKSSSTIAASDILRNSPLEK---PISNAKSHPIPKSKFDFLSKKKEEBERK 125
QY 90 KTDLEE--IKVNISEFGLMRYATKMDILINVESTICAAIAS-----T 130
D 126 NKEKEKASVLPVSPFALPFAPILETIANVLGLVLAAGSQOPLMTLIFGRILTSFT 185
QY 131 FQRLMLQIST-----DEFYDELKRVLYFYVLIIGFVYVYSTVGTFTY 175
D 186 NYAVIANQISOGGLTPETSALQAADLKQSGHMLALYLAIGMELATWLYMTFNV 245
QY 176 TGEHATOKIRRYLESTIRONIGYFDKAGEVTRITADNLIDGISEKVGILTLMA 235
D 246 TGLNSKRIRRYTAAVLRQELAFDDGAEVATRIOTDHLVOGSEKVALVPOYAG 305
QY 236 TFWTAFTIAYVYKWKALICSTIVATLVLTGSGSOPFIKYSKSLDSYGAGTVAEEVI 295
D 306 TFWGQVFLAFVRSPLRIGALVSLIPVIMLGGIMMTAKGTALDHIAKAGSLAEVI 365
QY 296 SSINATAFGTQDLAKOYEVHLDAAEKWGTKNQIWMGFMIGAMPGLMYSVGLGFWGS 355
D 366 GSIRTVQAFGEKILGDFAHDHIESKIVGRKSGIFEGFGLSIMEFYIYAVALAFYGG 425
QY 356 RFLVDGAVDVDDILVYLAALIGSPSLGNSVPMQAFNNAVAAAKIIGCTIDROSPIDPY 415
D 426 ILVSGAODSIVINVFSSILIGSSMMALBELAAVAKAGAAKLPATIDRPAIDSA 485
QY 416 SNEKRTIDFEGHIELRNKHIPESEVYMEVDSLSMPAKTTALVPGSGSKSYVG 475
D 486 SEEGFKPGLGELSEFENVKFHPSPRPILPKGTTTFTEAGKTFALVAGSGSGSYVS 545
QY 476 LVKEFPMYRQVTLDDGDIDMLNRILROQISLVSEQVYFPTTIYNHGLIGTYE 535
D 546 LIERFYPVSGVYKLDGDIRSLNMLRQIIGVSEPTLFGTYVRNVHGLIGSYE 605
QY 536 NESDEKVELLENKAKMANADFTALPEGETVNGGFFLSSGQKORAIARAVSDP 595
D 606 NASLEKFEVLYKACVADANANFIMKLPQGDITVNGEGMILSGGQKORAIARAVSDP 665
QY 596 KILLDEATSLDTRKSEGVQALERAEGRTTIVIAHRLSTIKTANIVLVNGKIAEQ 655
D 666 RILLDEATSLDTRKSEGVQALERAEGRTTIVIAHRLSTIKTANIVLVNGKIAEQ 725
QY 656 GTHDELV--DRGATRKLYEADRIEQRADALE--DAADIEDLTNADIAKTAASASDLD 713
D 726 GSHNDLLENNGPYAQLVNNOKLQEAALAEALQYDDIEDPDA-----VFIGSSSPMQ 779
QY 714 GKPTTIDRTGTHKSVSA-----ILSKRPETT-----PKYSLMTIL-----KFVASF 756
D 780 EKDQQLRAVYGRSLASIAMDIOAKRAEYVAGEKIPSSFGILARKLRMSADKFI--- 836
QY 757 NRPEIPYMLGLVPSVLAAGGQPOAVLYAKAISTLSPEQSKLRHADAFWSLMEFYV 816
D 837 -----YIAFIAICAGMYPSIALILFKALSDFEIQDP--AELRALRSRLMWYIT 887
QY 817 GIIDFITOSTGCAFAVCESEILIRARSTARTIIRQDIAFDXEENSTGLTSLSTET 876
D 888 ALAAAFVIFQSSAFESRAGWDLNGLVRKKLTARTLRHDIEMFERDNSTGAVTSLADOP 947
QY 877 KHLGVSQVTLGTILMTSTLGAATIALAIGWKIALVCISVNPVLLACGFYRYMLAQF 936
D 948 QKVGGLGPTLIGYVQSGATIGGCTIGLCPILALIGIACIFILVSGGIRLKVYVLK 1007
QY 937 QSRKLAVBESANFACERTSSIRVASLTRRDWEIYHAQDAGRTSLISVRSLSLY 996
D 1008 DQRMKILHAASAHLASERAGAVKTVASLTRKQVRIYSALAKAPMKLNFTSIKSQLF 1067
QY 997 ASSQALVFPCALCFWVGTLGHEDIFREFVCSSEILFGAQSAGTVESFADMGKAK 1056
D 1068 AASQGLTFCTIALVYIGALWIIDAKYSTASFTVNLIVASIQAGNVTFVDPASKAN 1127

[illegible]

RESULT 4

S55692 multidrug resistance protein homolog (mdr) - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 27-Oct-1995 #sequence_rev:10n 03-Nov-1995 #text_change 02-Feb-2001
C:Accession: S55692
R:Castillo, G.; Shen, H.J.; Horvitz, S.B.
Biochim. Biophys. Acta 1262, 113-123, 1995
A:Title: A homologue of the mammalian multidrug resistance gene (mdr) is functionally ex
A:Reference number: S55692; MUID:95322451; PMID:759185
A:Accession: S55692
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1287 <CDS>
A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
F:420-614/Domain: ATP-binding cassette homology <ABC1>
F:437-444/Region: nucleotide-binding motif A (P-loop)
F:1062-1258/Domain: ATP-binding cassette homology <ABC2>
F:1079-1086/Region: nucleotide-binding motif A (P-loop)

Query Match	37.18;	Score 2493.5;	DB 2;	Length 1287;
Best Local Similarity	42.68;	Pred. No. 8.8e-143;		
Matches 544;	Conservative 215;	Mismatches 456;	Indels 63;	Gaps 12

QY	90	KTOLEEIIVNISPFGMLIAKRMJLLIMVSTICAIANAASFOKIMLY-----	127
Dd	39	KKEKTERPKAVGVMTMRYISTSPKMLMFEJTIASLHGALPLMLVPEGBMDSFYVNG	98
QY	138	QISYDF-----YDELTKNVLXYEYLIGEFVTVYVSTGVFIYGEHATQIR	185
Dd	99	QVDFGNFWMESMINASRELQGMHTTAYVYSGLOFGVMLCAVIOISWTLISAGHQIKIR	158
QY	186	EYTESIIRONIGYFDKLGAGEVTTTRITADPTNLQDSISEKVGITLTALATFYVATIIAY	245
Dd	159	SNFHAVYLRQIEGMFNDINDAGELNTRPLTDVSKINEIGDKIAMLLQSTLTVLTVGFIIGE	218
QY	246	VKYMKLALICESTIVALVLMVGGSSQRLIKYSKSLDSYGAGGVVAEEVJSSIRNATAFG	305
Dd	219	IKGRKLTVMWGAISPIINGLSAIIWAKVLSAFTNKEIKAVAKAGVAEEVJSSIRTYEAFG	278
QY	306	TODKLAKOYEVLHDEAEKMGTKNOIVGPMIGAMFGMTSNYGSGFMGSRPLVDGAVDV	365
Dd	279	GQNEKIHRHYENLMDAKKIGIKKAKITLVNVSIGAFELMTIAYVSLAFYGTLLIIDGGYTI	338
QY	366	GDLITVLMALIGSFSLGANSVPAQAFTTNVAAAALFEFTIDROSPLDPSYNEGKTTDHF	425
Dd	339	GSVLTVEFAVYIIGAFVAGQSPNTEAFVANNRGAAYITFNIIIDNPKIDRSKRGKLPDKI	398
QY	426	EGHIELNRKHIIYSRDEVTVMEDVSLSPMAGKTTALVPGSGSGKSTVGVLEFRTYMPAR	485

Db	399	KGDIEFKNVITPYDSRKDIQVGLKGLNIMPSSGKTVALVSSCGKSTYQVLIQRRIYD	458
Qy	466	GTVLDDGHDIDYDLRLRLROISVSGEPVLTFTYTKINIRGLIGTKENESDKREL	545
Db	459	GVYTLDDGDRSLNIRILRELTIGVSGEPILFTPTTJADNIRG-----REDVTKEE	509
Qy	546	IENAKMANAHDFYALPEGYETNVCORGLISGGOKRIATARAIVSPDKILLDEATS	605
Db	510	IERTAKENAMVDLMKLPDKLLETUVGERGQLSGGOKRIATARLVRPKILLDEATS	569
Qy	606	ALDTKSEGVQALERAEGRTTVIAHRLSTTKTANITVVLNGKIAEGDTHDELVRG	665
Db	570	ALDTSESEVQALDKAREERTTVIAHRLSTIRANALAGPDNGIVGEGSHKELMERG	629
Qy	666	GAYRKIVDAQINQRADLEADAE-----DLTNADIAKTAASSSDLDOKPPTI	719
Db	630	GVYFNLVTLQVETSKOTE--EDLEHNIYEKKIPVHTSHNLVRRKSSMNTIKSVPETE	687
Qy	720	DRTHGKHSVSSALITSRPPTTPKYSIMTLTKFVASFNPBRLEPYMLIGTVSVLAGGQP	779
Db	668	D-----KEYDDE--EKKKEGGPPVSPFKMK-----LNPMPYRYVGVICAMIGATOP	736
Qy	760	TOAVLYAKAISTSLPESQYSKLRHADPWSLMFVYVGIOTOSTNGAAFAVCSERLI	839
Db	737	AFALIFSRILIGVFAGPYQ--MRSESSMYSILFIALGVSEFTPEFLQRTFGKAGEILT	793
Qy	840	BRASVARTLITRDIAFEPDKENSNUGALTSPLSTETKILSGVSVLTGLTILMSTYIGA	899
Db	794	MRLRGSKSMRLQELIEMFDSKNSGALTTRLATDASOVGATETRIALLAQVYANLGT	853
Qy	900	AIITIALIGMKLAVVCISVVPVLLAGGFYRFYLMQFOSRKLAEGSANGFACETSIR	959
Db	854	AIITISFYGMQLTLLIATAYPVTAAGIVEMKMFAGHAKKDKKELEKNGKISTDAVINR	913
Qy	960	TVASITREBDWEYETHAQDLAQRTSLISVLSLTYASSQALVFCCYALGFWTGTLL-	1018
Db	914	TVVASITREKFEAYETKSLSGPYRNBKIRKAHLGLTYGLSQAHVYLCJCFWFSYLVGAVLY	973
Qy	1019	--GHEHDVIFPEFVCFSEILTFQSGASGTYFSPADMGKAKNAAEFRRLPDKRPIDOWS	1076
Db	974	VEGLMKID--EVLVYSALYVGLMAGLCOTSSRPADYTKAMISAHTFSLIERVPOIDYS	1031
Qy	1077	EEGKLETVGEELFEFNNVHERPYTRPEQVLRGLDLYVRGQYVALVPSGCGKSTTIAL	1136
Db	1032	DQGEKPKNGSGNVYFKVNNYPTRPDIYVGLDIDISYQGETIALVSGSGKSTYVSL	1091
Qy	1137	LEREYVDIAGSIIVDGDIDISKTINSTRPSLISVSEPTLYOGCTIKENILGLIYEDVPE	1196
Db	1092	LEREYDPEEGVILVDGSIYKNLNIQWRVRAQMGIVSEPTLTFPCSIGDNIAGDNNRKVTO	1151
Qy	1197	EFLTKACKDANIYDFIMSLPEGEPTVYVSGKGLSGGOKORVAILARALLRPPKILLDEA	1256
Db	1152	BEITAKENINHSFIESLTDKNTVNGDKGQVLSGGOKRIATARLVRPKILLDEA	1211
Qy	1257	TSALDSESEKVOQAALDAAARGFTTIAVAHRLSTIQAADVIYVFDQKIVESGTHSELVQ	1316
Db	1212	TSADTSESEKVOEALDKAMGRGTCIVIAHRLSTIONADKIAVIONGRVYEGTHQOULO	1271
Qy	1317	KKGRVYELVNLQSIGKGH 1334	
Db	1272	LKGVYFSLVTVIQ--LGH 1286	

RESULT 5

DWH01 multidrug resistance protein 1 - human
 N/Alternate names: P-glycoprotein 1
 C/Species: Homo sapiens (man)
 C/Date: 31-Dec-1990 #sequence, revision 18-Aug-1995 #text, change 19-Jan-2001
 C/Accession: A34914, P50162; S15050; A25059; S43836; I52238; I65204
 R/Chen, C.; Clair, D.; Ueda, K.; Pastan, I.; Gottesman, M.M.; Roninson, I.B.
 J. Biol. Chem. 265, 506-514, 1990

A:Title: Genomic organization of the human multidrug resistance (MDR1) gene and origin
A:Reference number: A34914; MUID:90094448; PMID:1967175
A:Accession: A34914
A:Molecule type: DNA
A:Residues: 1-1280 <CHE>
A:Cross-references: GB:M29447; GB:J05168; NID:g187496; PIDN:AAA59576.1; PID:g386862
R:KioKa, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to JIPID, April 1991
A:Reference number: PS0162
A:Accession: PS0162
A:Molecule type: DNA
A:Residues: 1-22 <KIO>
R:KioKa, N.; Yamano, Y.; Komano, T.; Ueda, K.
submitted to the EMBL Data Library, April 1991
A:Description: Transcriptional regulation of multidrug resistance gene (MDR1) expression
A:Reference number: S15500
A:Accession: S15500
A:Molecule type: DNA
A:Residues: 1-22, 'R' <K12>
A:Cross-references: EMBL:X58723; NID:g34522; PIDN:CAA41558.1; PID:g34523
R:Chen, C.; Chin, J.E.; Ueda, K.; Clack, D.P.; Pastan, I.; Gottesman, M.M.; Roninson, I
Cell 47, 381-389, 1986
A:Title: Internal duplication and homology with bacterial transport proteins in the mdrl
A:Reference number: A25059; MUID:87028230; PMID:2876781
A:Accession: A25059
A:Molecule type: mRNA
A:Residues: 1-184, 'V', 186-1280 <CH2>
A:Cross-references: GB:M44758; NID:g187468; PIDN:AAA59575.1; PID:g307180
R:Chambers, T.C.; Pohl, J.; Glass, D.B.; Kuo, J.F.
Biochem. J. 299, 309-315, 1994
A:Title: Phosphorylation by protein kinase C and cyclic AMP-dependent protein kinase of
A:Reference number: S43838; MUID:94220047; PMID:7909431
A:Accession: S43838
A:Molecule type: protein
A:Residues: 656-689 <CHA>
R:Geckeler, V.; Weger, S.; Probst, H.
Biochem. Biophys. Res. Commun. 169, 796-802, 1990
A:Title: mdrl/P-glycoprotein gene segments analyzed from various human leukemic cell lines
A:Reference number: I52238; MUID:90290529; PMID:1972623
A:Accession: I52238
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 178-215 <RES>
A:Cross-references: GB:M37724; NID:g183537; PIDN:AAA88047.1; PID:g553314
A:Accession: I65204
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 800-856 <RE2>
A:Cross-references: GB:M7725; NID:g183538; PIDN:AAA88048.1; PID:g553315
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell lines
functionally and functionally unrelated lipophilic antitumor drugs.
C:Genetics:
A:Gene: GDB:PGY1; MDR1
A:Cross-references: GDB:I20712; OMIM:171050
A:Map position: 7g21-7q21
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; phosphoprotein;
F:1-638, 653-1280/Region: duplication
F:49-350/Domain: hydrophobic <HB1>
F:351-637/Domain: hydrophobic <HB1>
F:410-604/Domain: ATP-binding cassette homology <ABC1>
F:4427-444/Region: nucleotide-binding motif A (P-loop)
F:551-555/Region: nucleotide-binding motif B
F:638-708/Domain: linker <LIN>
F:709-993/Domain: hydrophobic <HB2>
F:994-1280/Domain: hydrophilic <HL2>
F:1053-1249/Domain: ATP-binding homology <ABC2>
F:1070-1077/Region: nucleotide-binding motif A (P-loop)
F:1196-1200/Region: nucleotide-binding motif B
F:91, 94, 99/Binding site: carbonylate (Asn) (covalent) #status predicted
F:433/Binding site: ATP (Lys) #status predicted
F:661, 667, 671/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted
F:667, 671, 683/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status predicted

E:1076/Binding site: ATP (Lys) #status predicted
 Query Match 37.0%; Score 2489; DB 1; Length 1280;
 Best Local Similarity 42.1%; Pred. No.1.6e-142;
 Matches 535; Conservative 229; Mismatches 448; Indels 60; Gaps 9;

OY	94	ELIKNISPFGLMRATKMDLLIMVSTICLAIAAASFORIMLY-----	137
DB	28	KKKPTVSFVSFRRSSNMCLKLYMVGTLAAIIHGAGLPLMLVGEMLDIPAMGNLED	87
OY	138	-----QISDEEY-----DELTKNVLVFFYLIGSEFVTVYVSGVITYGEHATOKI	184
DB	88	LMSNTNSDINDIDGFEFMNLEEDMTIRYATYIGIGAVLYAAIYOVSTFCLAAGROIKI	147
OY	185	REYLESLIRONIGYFDKILAGEVTTRTIADTNLIODGISEKVGITLTALATVTAFLIA	244
DB	148	RKQFHAIRMGRIQGFVFDVHDVGLMTRLTDVSKINEGIDKRIQGFQSMATFFGIYG	207
OY	245	YKRYKALIGSITVALVLTMGSSQFIRIKSKSDSYGAGGYAAEEVISIRNATAF	304
DB	208	FTRGKRLVLVLAISPVGLSAVAWAKLISSTPDKELLAYAKAGVAEEVLAIRTVIAF	267
OY	305	GTPDOLAKOYEHLDEAEKMGTKNOIVMGFMGLMYNMGVGAFFMNGSRFLVDGAND	364
DB	268	GQKKELERIKNNLEAKRIGIKAKITANISIGAAFLLIYASTALAFWGTTLVLSGEYS	327
OY	365	VGDILTVMALIGSFSLGAVNSPNAQAFNVAAMAAKIFGTIDROSPLDPYSENEKTLDH	424
DB	328	IGQVLTFFSVLIGAFVSGQAPSJLEAFANANGAAVEIFKIIDNKRPSIDYSKSGHKPDN	387
OY	425	PEGHLELRNKHITYSRPEVTVMEDVYSMPAGKTTALVPBSGSGKSTVGLVERPMV	484
DB	388	IKGNLEFRVHVSYSRKRKVKILKGLNLTKVQSGQVVALVGNSGCKSTVQMLQRLXPT	447
OY	485	RGTVLDDHGDIDMLNRLRQOISVSGOPVFGTTIKNINHGILGTFENESHDYRE	544
DB	448	EGMVSVDGODIKTIINVRFLREITIGVSDPVLPAITIAENIYG-----REWVTMDE---	499
OY	545	LIENAKAMANAHDFTTALPEGETYVWVGQGFLLSGGOKORIAIARAVSDPKILLDEAT	604
DB	500	-IEKAVKEANAVDFIMKLPKHFDTLVGERGAQLDSGGQKORIAIARALVANNPKRIILLDEAT	558
OY	605	SALDTRKSEGVOALERAEGRTTVIAHRSTITTAHNIYLVNGKIAEGOTHDELVDYR	664
DB	559	SALDSESEAVVOVALDARKGRITTVIAHRSTVARNADYIAFGDGVITYEKNHDELMAKE	618
OY	665	GGAYKRLVEAORINOKREADLEDADIEDLNADIAKIKRTASASDDLDGRKTTIDRGCT	724
DB	619	KGIYKTLVMTQAGNEVE---LENA-----ADSKSKSIDLEMSNSDRSLIRKNST	668
OY	725	HKVSASAILSKRP-----PETTKRYSIMTLTKFEVASHNREIPYMLIGLFFSVLAGGG	777
DB	669	RSRVSGQAQDRKSLSTKEALDESIPYPSWRMK---LNLIEWYFVYGVCAIINGEL	724
OY	778	OPTAVLYAKAISTLSLPSQYSKLRHDADEWLMFPVVGIIQITOSTNGAAPVCSER	837
DB	725	QPAFIITFSKIIGVTRIDDPETK-RONSNTLSLFLAIGISFTFFLGQTFPGTKAGEI	783
OY	838	LIRRSARSTAFRTILKODIAFPDENKSNAGLISFLSTETKHLGSVSGVTLGIIMTSTLL	897
DB	784	LTKRLRYVFRSMLKODVSWEDDPKNTTGALTTRLANDAAOVKGAIGSLAVITONIANL	843
OY	898	GAAILTALAIGKTLAVICISVVPVLLAGGFYRFYMLAOFOSRSKLAYESANFACEATSS	957
DB	844	GGIITISFYMQQLLLLLAIVPIAIIANGVEMKRLSGALKDKKDEBAGAIATAEIEN	903
OY	958	IRTVASLTREPDWEIYHAQLDAGRTSLISVLRSSLLYASSQALVFPCVALGFWGGTL	1017
DB	904	FRTVSLTIOEQFTEHMYAQSLOVPRNSLRKAHIGITFSFGQAMMYASYACFFRGAVL	963
OY	1018	LGHHHYDIFRRFYVCSLILFGQASGTYFSFAPDMGKAKNMAAEFRRLDEDRKPOIDNNS	1077
DB	964	VAHKIMSEDEVLLVVSAYVFGAAGVQSSFPADPAKAKKISAAHIIIMITEKPLIDYST	1023

QY 843 RSTAFPTTIRDIAPFDKBEENSTGALTSTLSTFKHLSGVSTGLTILMTSTTLGAAIT 902
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 786 RWVWFSKMRDVSFMDPNKNTGALTTRLANDAGVKGATGARLVITTONINANTGII 845
 QY 903 IALAIKMLALVCISSVVPVLLACGFYRFYLAQFQSRKLAEGSANFACNATSIRTV 962
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 846 ISLIYQWOLTLTLLALVPIIALAGVEMKMLSGOALKDKKELGSGKIAEALENFTV 905
 QY 963 SLTRRDVWEIYHQAOLDAQGRSTLSILVRSLLYASSOALVFECVALGFEVGGTLLGHE 1022
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 906 SLTRBCKFENMVAQSLQIYRNALKKAHVGFIFSTQAMMYTSYAACFGAYLAREL 965
 QY 1023 YDIERFVCFSEILFQAQAGVTFSPADMGKAKNAAEFRRLDPRKQIDNMSEGEKL 1082
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 966 KTFEENLVLSAIVFGAMAVGVSSAPDPAKAKVASHITIMEKVPISIDYSTGLKP 1025
 QY 1083 EHYEGRIERNVHFRTPEQPVLRGLDLYTPRQGVYALVGPSSGCKSTTIALLEFYD 1142
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 1026 NTLGKVNKFEVFNPTPRDIPVLOGLNEVKKGOTLAVGSSGCKSTVVOLEEFYD 1085
 QY 1143 AIAGSLVVDGKDISKLINSYRSFLSLVSOEPTLYOGTIKENITLIGIEDVPEEFILKA 1202
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 1086 PMAGTFLDQKEVNOQVOMLRHLGIVSEPTLFDOSTIENLAKGDNRYVSGDETERA 1145
 QY 1203 CNDANTYDFIMSLEPGFNTVYVSGKGMLSGGQKQRYAIALRLDPRKILLDETSALDS 1262
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 1146 AKRANHOFTESLPDKYNTFRVGDGTQLSGCKQRIAIARALVQPHILLDEATSAIDT 1205
 QY 1263 ESEKVVQALDAAARGRTTAVAHRLSTQKADVIYVFDGKIVSESTHSELVQKGRY 1322
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 1206 ESEKVVQALDAAARGRTTAVAHRLSTQKADVIYVFDGKIVSESTHSELVQKGRY 1265
 QY 1323 ELYNLSLGR 1332
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 1266 SMVSVOAGAK 1275

RESULT 7

A34786 multidrug resistance protein 1a - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 02-Feb-2001
 C:Accession: A34786; A35671
 R:Devault, A.; Gros, P.
 Mol. Cell. Biol. 10, 1652-1663, 1990

A:Title: Two members of the mouse mdr gene family confer multidrug resistance with over-
 A:Reference number: A34786; MUID:90205845; PMID:1969610
 A:Accession: A34786
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1276 <DEV>

A:Cross-references: GB:M30697; NID:919911; PIDN:AAA39517.1; PID:9387429
 R:Hu, S.I.H., Cohen, D., Kirschner, L.S., Lohstein, L., Hartstein, M.; Horvitz, S.B.
 Mol. Cell. Biol. 10, 3596-3606, 1990
 A:Title: Structural analysis of the mouse mdr1a (P-glycoprotein) promoter reveals the b

A:Reference number: A35671; MUID:90287150; PMID:1972547
 A:Accession: A35671
 A:Status: preliminary
 A:Molecule type: mRNA

A:Residues: 1-525, 'HV', 528-938, 'S', 940-1035, 'V', 1037-1276 <HSU>
 A:Cross-references: GB:M33581; NID:9199104; PIDN:AAA39514.1; PID:9387427
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
 C:Keywords: ATP; duplication; glycoprotein; membrane protein; nucleotide binding; p-loop

F:423-431/Region: nucleotide-binding motif A (P-loop)
 F:547-551/Region: nucleotide-binding motif B
 F:1049-1245/Domain: ATP-binding cassette homology <ABC2>
 F:1066-1074/Region: nucleotide-binding motif A (P-loop)
 F:1192-1196/Region: nucleotide-binding motif B

F:429/Binding site: ATP (Lys) #status predicted
 F:1072/Binding site: ATP (Lys) #status predicted

Query Match 36.8%; Score 2471.5; DB 2; Length 1276;

Best Local Similarity 41.4%; Pred. No. 1,9e-141;
 Matches 527; Conservative 228; Mismatches 468; Indels 49; Gaps 6;
 QY 90 KTOLEEKVNISSFGILRRVATKDDILIMYSTICATAAATFORIMY----- 137
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 23 KKEKKKPAVSVLIMRIRYGLDRLYMLVGLTAAIHGVALPLMLIPEDMDTSFASVG 82
 QY 138 -----QISDFEYDELTKNVLYFYVLIGIEFVTVSVTGFITYGENATQK 183
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 83 NVSKNSTNMSEADKRAFALTEEMTYAYYIGIGVLYVATVQVSFNCIAAGRIHK 142
 QY 184 IREYVESILRQNIYFDKLAGEVTRITADYNLIDGISEKVLGTLTALAFVTAFT 243
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 143 IROKFHAIMOEGMDVDVDELNTRLDVDSKINEGIGDKIGMFFQAMATFFGCFII 202
 QY 244 AAYKYMALICSTYVALVLTMGSSQFIKYSKSLDSYGAGTAAEEVSSIRATA 303
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 203 GFTRGWKLTVLILAIISPVILISAGIYMAKILISFDKELHAYAKAGAAVEVLAIRTVIA 262
 QY 304 FGTQDKLAKQVEHLDEAEKMGFTKQIYMGFMIGAMGLVSNYGLGFMNGSRPLVDGAV 363
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 263 FGQKKLELEYNNNLEAKLGITKATTANISMAANLLIYASVALAFWGTSLVISEK 322
 QY 364 DVGDILTVMAILLGSFSLGNVSPNAQFTNAAVAAAKIFGTIDRQSPIDPYSNEGTLD 423
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 323 SIGQVLTVEFSVLIGAVSQASPNIEAFANARGAAYEVKIIDNKPSIDSFSGCHKPD 382
 QY 424 HFEGHIELRNVKHYPRPPTVMEVDSLMPGKTLALGPGSSGSGTYGVLYERYMP 483
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 383 NIOGNEFKNIHSYPRKEVQILKGLNLVKSQVTLVAGNSCGKSTVVOLEQRLYDP 442
 QY 484 VRGTVLIDGHDIKDLNRLRQOISLVSOEPTLYGTYIRNRHGLGTYENSEDEKVR 543
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 443 LDGWSLIDGDIRINRYLREIIGVVSQEPVLEATITAEIRKG-----REVTM 493
 QY 544 ELIENAAKMANHDETTALPEGETTNAVQGFELSGGQKQRIARAIVSDPKILLIDEA 603
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 494 DEIEKAVKANAAYDFIKLHPQFTLVGERGAOLSGGOKQRIARALVNPKILLIDEA 553
 QY 604 TSLDTESEGVQALERRAEGRTTIVAHRLSTIKTAHNIVYLVNKKIAEQGHLEVD 663
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 554 TSLDTESEGVQALERRAEGRTTIVAHRLSTIKTAHNIVYLVNKKIAEQGHLEVD 613
 QY 664 RGAAYRKLVEAQRINDEKADALDADADLTNADIKITASASASDLDGKPTTIDRTG 723
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 614 EKGTFELVMTQIRGNIELGNCKSKDELNDLMS---SKDSGLIRRSRKSICG 670
 QY 724 THRSVSSAILSKRPETTPKYSILMTLLKFAVAFNRPEIPYMLIGLVSVLAGGQPTQAV 783
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 671 PHQDRKLTKEALDEDDVPAPSEFRILK-----LNSTEMPYFVVGIFCAILNGIGLPASFV 726
 QY 784 LYAKAITLS---LPESQYKLRHADPWSLMPFVVGIIQFTIQTSTGAAFAVSEKLIR 840
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 727 IFSKVVGVFTNGGPPETQ---RONSMLFSLFLILGIIISFTFFLOGFTFGKAGELITLK 782
 QY 841 RARSTARPTLRDIAFDKEENSTGALTSTLSTFKHLSGVSTGLTILMTSTTLGAA 900
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 783 RLKRYMFKSMRLDVSFMDPNKNTGALTTRLANDAGVKGATGARLVITTONINANTGII 842
 QY 901 IIALAIKMLALVCISSVVPVLLACGFYRFYLAQFQSRKLAEGSANFACNATSIRTV 960
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 843 IISLIYQWOLTLTLLALVPIIALAGVEMKMLSGOALKDKKELGSGKIAEALENFTV 902
 QY 961 VASITRRDWEIYHQAOLDAQGRSTLSILVRSLLYASSOALVFECVALGFEVGGTLLG 1020
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 903 VVSLTRQKQETMAQSLQIPYRNAMKKAHVFGITTFPTQAMMYTSYAACFGAYLAREL 962
 QY 1021 HEVDIFPFVCFSEILFQAQAGVTFSPADMGKAKNAAEFRRLDPRKQIDNMSEGEKL 1080
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 Db 963 QLMTFEENLVLSAIVFGAMAVGVSSAPDPAKAKVASHITIMEKVPISIDYSTGLKP 1022
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|
 QY 1081 KLETFEGERIERNVHFRTPEQPVLRGLDLYTPRQGVYALVGPSSGCKSTTIALLEFYD 1140
 |:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|:::|||||:::|

Db 1023 KPNMTEGNVQSGFVFNPTFRSIPVLOGSLSEVKKGQTLALVSSGCKSTVOLLERF 1082
 1141 YDAIAGSILVQDKDISKLNINSRSLVSOEPTLYOGTIEKINILLGIVEDVEEPLI 1200
 1083 YDPMAGSVFLDKEIKLNOLVOMLRQOLGIVSOEPLIFDQSIENIATVAGNSVSEELIV 1142
 Oy 1201 KACRANAIYDFMISPEGNIVVSGKGMLSGGOKORVAAIAALRLDPKILLIDEATSL 1260
 Db 1143 RAKRANIHQFIDSLPDKRNTVCGKGTQOLSGOKORAIARALVQPHILLIDEATSL 1202
 Oy 1261 DSESEKVVQALDAARGRHTTAVAHRLSTIQKADIVYFQDKIVESGTHSELVOKKR 1320
 Db 1203 DRESKVVQOALDKRAREGRTCIYIAHRLSTIQNADLIVYIQNGKKEHSTHQDLAKGI 1262
 Oy 1321 YVELVNLQSLGR 1332
 Db 1263 YFSMVSVQAGAK 1274
 RESULT 8
 DVM51
 multidrug resistance protein 1 - mouse
 N:Alternate names: P-glycoprotein 1
 C:Species: Mus musculus (house mouse)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
 C:Accession: A33719; A25057; I57510
 R:Raymond, M.; Gros, P.
 Proc. Natl. Acad. Sci. U.S.A. 86, 6488-6492, 1989
 A:Title: Mammalian multidrug resistance gene: correlation of exon organization with stru
 A:Reference number: A33719; MUID:89367274; PMID:2570420
 A:Accession: A33719
 A:Molecule type: DNA
 A:Residues: 1-1276 <RAY>
 R:Gros, P.; Croop, J.; Housman, D.
 Cell 47, 371-380, 1986
 A:Title: Mammalian multidrug resistance gene: complete cDNA sequence indicates strong ho
 A:Reference number: A25057; MUID:87028229; PMID:3768958
 A:Accession: A25057
 A:Molecule type: mRNA
 A:Residues: 1-1276 <GRO>
 A:Cross-references: GB:M14757; NID:g199100; PIDN:AAA39513.1; PID:g387426
 R:Raymond, M.; Gros, P.
 Mol. Cell. Biol. 10, 6036-6040, 1990
 A:Title: Cell-specific activity of cis-acting regulatory elements in the promoter of the
 A:Reference number: I57510; MUID:91042535; PMID:2248681
 A:Accession: I57510
 A:Status: Preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-21 <RES>
 A:Cross-references: GB:M60348; NID:g199102; PIDN:AAA39513.1; PID:g554199
 C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
 C:Genetics: structurally and functionally unrelated lipophilic antitumor drugs.
 A:Gene: mdr1 (Pgp1)
 A:Introns: 21/2; 38/3; 92/1; 111/2; 175/2; 233/3; 274/2; 332/3; 370/3; 407/3; 449/3; 517
 C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
 C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr
 F:1-637; 653-1276/Region: duplication
 F:409-637/Domain: ATP-binding cassette homology <ABC1>
 F:428-433/Region: nucleotide-binding motif A (P-loop)
 F:550-554/Region: nucleotide-binding motif B
 F:1051-1247/Domain: ATP-binding cassette homology <ABC2>
 F:1068-1075/Region: nucleotide-binding motif A (P-loop)
 F:1194-1198/Region: nucleotide-binding motif B
 F:73_91_96_103/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:433/Binding site: ATP (Lys) #status predicted
 F:1074/Binding site: ATP (Lys) #status predicted

Db 23 KKEKREKPAVGFVGFRRADWLDKICMLTGTLAAIHGTLPLMLTVGNMTDSFTKAE 82
 137 -----YOISYDEFEDELKKNLYFEYLIGEFVYVSGEYITGEHA 180
 Oy 83 ASILSITNOSPNSITLITNSLSLEEMAIYAYITIGAGVLAIVYIVSILCAGAQ 142
 Db 181 TOKIREYLESILRONIGYFDRKLAGEVYTRITADTNLIDGISEKVGTLTALATFYTA 240
 Oy 143 IKHROKFFPHALIMNOEIGFVDHVGELNTRLDVSKINDIGIKGFQSIPTPLAG 202
 Oy 241 FIITAYKYKMLLCSSITVALVLMGGSGQFIITKSKSLDSYGAGTVAEEVSSITN 300
 Db 203 FIIGFISGKRLVTLVLAASPLGLSALMAKVLSTNLEQAYAKAGVAEEVLAAT 262
 Oy 301 ATAFGTQDKLAKQYEHVHDEAEKMGTKNOIVMGFMIGAMEGLMYSNYGCFPMGSRFLD 360
 Db 263 VAFAGQCKELEERKNKNELEAKNVGIKKAITSIGIVLYIVASYALAFYGTSLYS 322
 Oy 361 GAVDVGDILTVLMAILIGSFLGNVSPNAQFTNVAATAKIFGTIDRQSPDPYNECK 420
 Db 323 NEYSIGEVLTVFESILGTFSIGHLAPNIEAFANRGAFAELFKILLDNPSIDSTKGY 382
 Oy 421 TLDHEGHELANVKHIYPSREYVYMEVYSMAAGKTTALVGPSSGSKSTVGLVERF 480
 Db 383 KPDSTMGNLFEKNVFNHFNPSRSEVOILKGLNLRKVSQGTVALVGNSSGCKSTVOLMORL 442
 Oy 481 YMPVRGTVLLDDHDKDLRWLRQOISLVSOEPLVFGTIIYKKNRHGLIGTKYENESD 540
 Db 443 YPDEGVYSIDQDRTIVRKLRELIGVSOEPLFATTIENIRYG-----RED 493
 Oy 541 KYRELLENAAKANAHDFTALPEGETNVGQGFLLSGGOKORIAARAVSDPKILL 600
 Db 494 YTMDEIEKAVKANAADFIMKLPDQFLVGBRQALSGGOKORIAARAVRNPKILL 553
 Oy 601 DATSALDQKSSGVQAALEERAEGRITVIAHRLSTITAINIYLVNKGTAEGTDE 660
 Db 554 DEATSLDTEESAIVQOALDKRAREGRTVIAHRLSTVNAADVAGFDGIVTEGNDHE 613
 Oy 661 LVDRGAVYKLEVAQ-RINEQKAD---ALBEDAEDLTNADIKIKTASASASDLQK 715
 Db 614 LIREKGIYFKLMTQTRNGEIEPKNNAVGSODTASELTSSEKSPILRRIYYSVHK 673
 Oy 716 PPTIDRTGHNKSVSAIISKRPETTPKYSMTLTKFVASFNRPETPYMLIGLVSVLAG 775
 Db 674 QOERRLSMKEAVD-----EDVPLVSFWRII---NLMSWPYLLVGLVAING 720
 Oy 776 GGQPGVAVLYAKAISTLSPEQSGKLRHDADFMGLMFPVGIIOFIQSTGNAFAVCS 835
 Db 721 CIQPVFAIVFSRIVGVFSKDDDEHK-RQNCMLFSLFLVWGLISVYVFGGTFGKAG 779
 Oy 836 ERLIRARSTARFTLLRODIAFFDKEENSTGALTSFLSTFHLISGVSGVTGLTLMST 895
 Db 780 ELTRKVRNRVKSMLRODISWFDHKNSTGLTRIASDASVYGAMARLAVYQANA 839
 Oy 896 TLGAALITLAIKMLALVCISVPVYLLACGFRYRYMLAQFQSRKSLAYEGSANFACENT 955
 Db 840 NGGTGIVSLVYGQDITLLVVIIPLIYIGITLKKLISQALKRKKOLEISGKIATEAI 899
 Oy 956 SSIRVASTLRRRDWEIYHAOLDAGRTSLSVLRSSILVASSQALVFCVALFWRG 1015
 Db 900 ENKRTIVSLTRQOKETMETNAGSLQVPRNAMKKAHVGTGTFQAMMFSTAAQFRFA 959
 Oy 1016 TLGHHEDYDFFPFCSEILFGAQSAGVFSFADPMCKAKNAAEFRILPRKFOIDMW 1075
 Db 960 YLVAQOLMFENVMVLVSNAVGAAGAGTSSFADYAKAKASASIIIRIIEKTEIDISY 1019
 Oy 1076 SEEGKLETVGEIEFRNVHFRYPTRPEQPVLRGLDLYKQOYVALVGPSSGSKSTYA 1135
 Db 1020 SNEGKPLTLEGNVKNGVQFVYPRPNIPVLQGISLEVKKQTLALVSSGCKSTVQ 1079
 Oy 1136 TLERVDIAGSILVDKDISKLNINSRSLVSOEPTLYOGTIEKINILLGIVEDV 1195

Query Match 36.5%; Score 2453.5; DB 1; Length 1276;
 Best Local Similarity 40.9%; Pred. No. 2.3e-140;
 Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;
 Oy 90 KTQLEIKVNIISFGIMKRTAKMDILINIVISTICAIAASTQRIML----- 136

Db 1262 QKCIYFSW 1270

RESULT 10

DVMS2
multidrug resistance protein 2 - mouse
N:Alternate names: P-glycoprotein MDR2
C:Species: Mus musculus (house mouse)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: A30409; S70711
R:Gros, P.; Raymond, M.; Bell, J.; Hausman, D.
Mol. Cell. Biol. 8, 2770-2778, 1988

A>Title: Cloning and characterization of a second member of the mouse mdr gene family.
A:Reference number: A30409; MID:88302195; PMID:3405218
A:Accession: A30409

A:Molecule type: mRNA
A:Residues: 1-1276 <HSU>
A:Cross-references: GB:J03398; NID:g199109; PIDN:AAA39516.1; PID:g387428
R:Kirschner, L.S.
Nucleic Acids Res. 24, 2829-2834, 1996

A>Title: De novo generation of simple sequence during gene amplification.
A:Reference number: S70711; MID:96313253; PMID:8759018
A:Accession: S70711

A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 43-92 <KIR>
A:Cross-references: EMBL:U46839; NID:g1228142; PIDN:AAC52722.1; PID:g1228143
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant cell
structurally and functionally unrelated lipophilic antitumor drugs.
A:Genetics: mdr2

C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; transmembrane pr
F:1-637/653-1276/Region: duplication
F:409-603/Domain: ATP-binding cassette homology <ABC1>
F:426-433/Region: nucleotide-binding motif A (P-loop)
F:1049-1245/Domain: nucleotide-binding motif B
F:1067-1074/Region: ATP-binding cassette homology <ABC2>
F:1192-1196/Region: nucleotide-binding motif A (P-loop)
F:88_94/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:432/Binding site: ATP (Lys) #status predicted
F:1072/Binding site: ATP (Lys) #status predicted

Query Match 35.5%; Score 2387; DB 1; Length 1276;
Best Local Similarity 42.1%; Pred. No. 2.5e-136;
Matches 534; Conservative 212; Mismatches 458; Indels 64; Gaps 13;

Qy 97 KVN-1SFGLMRYATKMDLLIWIYITICAIIMASTFORIML-YQISYDEYD----- 146
Db 33 KVNLLGLTLFRISDWQDLFMFLGILMAIHSGSLPLMIVFGEMTDKFDVDTGNFSLP 92

Qy 147 -----ELTKNVLYEYVLIGIEFVYVSTGVFIYGEHATOKIRRYLLE 190
Db 93 VNFSLMLNPGRLIEEMKRIYAYISGLGCVLVAYIYVSEPTLAAGQIKIRKRFH 152

Qy 191 SLIRONIGYFDKLAGETVTRITADINLIQDISSEKVGILLALAFVTAFTIAYVYK 250
Db 153 ALIRQMGWFDIKGTETELNTRLDVSKISEGIDKVFQFOIAFFAGFIYGFIRGMK 212

Qy 251 LALICSSITVALVLMGSSOFTIKSKSDSYAGGVAAEYVSIINNAFQIDKL 310
Db 213 LTLVMAISPILGLSTVAVAKILSTFSDELAAYAKAGVAAEAPAIRTVIAFGQNK 272

Qy 311 AKOEVHLDIAEKMGTKNIYVGMFGAMFGLMYSNYGFMGSRFLVDGAVDGIIT 370
Db 273 LERYQKHELNAKKIGIKKRAISANISMGILFLIYISYALAFYGTSLVSKETITNMT 332

Qy 371 VMAILIGSFSLGANVSPNAAFTNVAANAATFETIDROSPLDPPYSNEKTLIDFEGHTE 430
Db 333 VFESLILIGAFSVGOAFCIDAFANAKGAAYVIEDIINNPKIDPSERGHKDNKGNLE 392

Qy 431 LKNVKKIYSPREYVMEVSLSPAGKTTALVPGSGSKSTVGLVEREYMPKSTVLL 490
Db 393 FSDVHSYSPRANIKIKLGNLKVKSQGVVALVGNSSGCKSTVQLLRDLPTEKIST 452

Qy 491 DGHIDKDLNRMRLQOISLYSOEPVLEFRTIYKINRHLIGTYENESSEKVELLENAA 550
Db 453 DGODIRNFVRCLEIIGVSOEPVLESTTIAENIRG-----RGVYTMDE-----TEKAV 503

Qy 551 KMANADFTALPEGETYVNGRGFLLSGGOKORIAIARAVSDPKTILLDEATSLDTR 610
Db 504 KEMNAVDFIMKLPQKPTDLVGDGAQSLSGGOKORIAIARLVNPKILLDEATSLDIE 563

Qy 611 SEGVVQALERRAAEGRTTIVIAHRLSTKTAHNIYVLVNGKIAEGSTHDELVDGAYRK 670
Db 564 SEAEVQAAALDKAREGRTTIVIAHRLSTIRNADVIAFGEDGVIEGSHSELMEKKEGIYER 623

Qy 671 LVQAORINEOKEADALDEADAEILTANADIKITASSASDDLCKPTTIDRTGTHSVSS 730
Db 624 LVNMQTAGSQILSEFEVELSDERKAAGDVAP-----NGMKARIFRNSTKSLKS 672

Qy 731 AILSKRPET-----TPKSLMTLKFVASFNRPELPYMLIGLVPSYLAGGQPTQAVY 784
Db 673 PHQNRIDEETNELDANVPVPSFLKVLK-----LNKTEPFFVGVTCALANGALQPAFSTI 728

Qy 785 YAKAISTLSIPESQYSLRHADDFMSLMFVVGIIQFTOSTNGAFAVCSERLIRRAS 844
Db 729 LSEMAIFG-PGDDAVK-DOCKNMFSLVGLGLVLSFEFFLGGTFEKGAEILTRRLS 786

Qy 845 TARTILRODIAFPDEENSTGALSTFSTETHLSGVSVTILGTLMTSTIGAAITIA 904
Db 787 MAKRAMLRQDMWFDDHKSTGALSTRLADAQVAGAGTALIALAQNANAGTGTIS 846

Qy 905 LAIGMKLALVCSIVSVVLLACGFYRPMYLAOFOSRKLAYEGSANFACENTSSIRTVAS 964
Db 847 FIVGMQDTLLLSLVPEFVIAVAGIVEKMKLAGNAKRDCKEMEAGKATATAEIRIVVSL 906

Qy 965 TRRDVWEIYHAQDQGRSLISVRLSSILYASSQALVFCYALQFWYGTLL--GHHE 1022
Db 907 TORRKEESMVEKTLHPYRNSVYKHAHIGTFTSIQAFMYFSYAGCFRFGSYLIVNGHR 966

Qy 1023 Y-DIFFEYCFSESLFGAOSAGTVFSEFAPDMGAKNAAEFRLFPKQIDOWMSEGRK 1081
Db 967 FKVY---ILVFSALIVGAAVAGHASSFADPYAKAKLSAALVSLFEPKPLIDYSTSEGLM 1023

Qy 1082 LETVEGEIEFRNRYFPTRPEQPVRLGLDLYVKQOYVALVPGSGCKSTTALIEREY 1141
Db 1024 PDKFEGSVTFNEVYFVPTPRANVPVLQGLSLVEKKQGTALVSSGCKSTVQLLIEREY 1083

Qy 1142 DALAGSLVDGDKDISKLNINSTRSEFLSYSOEPTLYQGITKENILIGIYEDVDPEEFLIK 1201
Db 1084 DPMAGSVLLDDGDAKRLNQWMLAQIGIVSOEPILEDCSIENAIYAGDNSRVVPHDEIYR 1143

Qy 1202 ACKDANIYDFINSLPEGFMTVVGSGKGM.LSGGOKOVIAIARALIRPKILLDEATSLD 1261
Db 1144 AKKANIHPIETLPQKTYTRVGDKTQLSGGOKORIAIARALIRPKILLDEATSLD 1203

Qy 1262 SESEKVVQALDAARGRTTIVAHRLSTORADYVVEFDGKIVSGTHSELVOKRGYR 1321
Db 1204 TSESEKVVQALDKARGRFCIYIAHRLSTIQNADLVIENGVKRGHGHQDLAQKGIY 1263

Qy 1322 YELVNIQS 1329
Db 1264 FSNVNTQA 1271

RESULT 11

DVHUS3
multidrug resistance protein 3 - human
N:Alternate names: P-glycoprotein MDR3
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: J50051; S01346; A42213; 138015

F:429-436/Region: nucleotide-binding motif A (P-loop)
F:1054-1250/Domain: ATP-binding cassette homology <ABC2>
F:1071-1078/Region: nucleotide-binding motif A (P-loop)

Query Match 35.1%; Score 2357; DB 2; Length 1281;
Best Local Similarity 41.3%; Pred. No. 1.7e-134;
Matches 526; Conservative 224; Mismatches 449; Indels 76; Gaps 13;

```

QY 97 KVN-ISEFGIMRYATKMDILIMVSTICATAAATFORIML-YQISYDEF----- 145
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 36 KVNLIIGLLEFRYSMDQDKLFMLIGTIMALAHSGLEPLMIVFEEMDKFVNNAAGNSLP 95
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 146 -----DELTKNNLYFYVLGIGFVTVVSVGVFYTGSEHATQKIREYLE 190
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 96 VNEFLSMINRGRLIEEMTRAYVYSGIGGVLAAYIOVSFMTLAGROKIKRIONFFH 155
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 191 SILRONGYEDKIGAGEVTRITADTNLIODGISEKYGLELTALATFVTAFLIAYVWK 250
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 156 AILRQEMGFDDIKTTELNLRLDDDISKISBIDKVMGFQAVATFPAGFVGFIRGMK 215
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 251 IALICSTIYALVLTMGSSQFIKYSKSIDSYAGAGTAAEYISSIRNATAFGTDKL 310
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 216 LTIYMAISPIILGSAVAMKILSTFSDKELAAAKAGAAVEALGAIIRYIAGGQNK 275
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 311 AKQEVHLDAEKWTKNQIVMGFMIGAMTSGTGLGFWGSRFLVDGADVDTLT 370
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 276 LEROKHLENAKKIGIKKAISSANISMGIAFLIYASALAFWYSTVTSKRYIGNAMT 335
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 371 VLAAILIGSRLGNVSNQAFTNAVAANKIGFTIDRSPIDPSYSEKGLDHFEEHIE 430
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 336 VFFSILIGASVGOAACCDIFANARGAAVYFDIIDNPKIDSFSEKGRKPKDINKMLD 395
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 431 LRNVKHIYPSRPETVMEVDSLSMPAGKTALVPGSGSKSTVVGLEFERYMPARGVLL 490
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 396 FSDVHFESYPRANKIKIKGLNKLKVOQGYALVGNSSGGRKTTIQLRLQRLDPREGTISI 455
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 491 DGHDIKLNLMRLRQOISLYSOEYVLEGTITTKNIRHGLGTKEENSEDKRELINAA 550
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 456 DGDIDIRNFNRVYLEIIGVVOSEFLVSTIAENIRYG-----RGNTMEIKKAV 506
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 551 KMANAHFETALPGEYENNGORGFLLSGGOKRIATARAVSPDKILLDEATSDTK 610
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 507 KEANAYEFINKLPKFTPLVGERAQISGGOKRIATARAVSPDKILLDEATSDTK 566
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 611 SEGVQALERAAEGRTTIVAHRLSTIKTAHNTVVLVNGKIAOGTHDELVDGAYRK 670
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 567 SEAEVQAALDKARGRRTIVAHRLSTVRNADVIAGFEDGVIVGSGHSELMQEGYVFK 626
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 671 LVE-----AQRINQKADALDEADLTNAOIAKIKTASSASDLDGKRTTIDRGTH 725
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 627 LVNMQTSGSOILSQEFVEELSEKKAADGMP-----NGMKSHTFRNSTK 670
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 726 KSVASA-----ILSKRPETTPKYSIMTLKLFVASFNPETPYMLIGLVFSVLGGG 777
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 671 KSLKSSRAHHRLDVEDLDANVPVPSFLKYLK-----LNTMEPYFVGVCAIVNCA 726
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 778 QPTQAVLYAKAISTLSPESQSKLRHDADPMLFVVGIIQETITOSTNGAAEAVCSER 837
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 727 QPALSIIISEMIAIFG-PGDDAVK-QQKCNLFSLVGLIGVLSFTFPLQGFTEGKAGEI 784
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 838 LIRABASTARTITLRODIAFPDKENSTGALTSFLSTETKHLGSGVGYTLTILMTSTTL 897
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 785 LITLRLSMAPFAMLRDMSWEDDKNSTGALSTRADRAVOGATGRLLALIONANTL 844
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 898 GAATITIALAIGMKLATVLCISVVPVLLACGFYRFYMLAOFQSRSLKAYGSANFCEATSS 957
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 845 GTGIIISFTYQMLTLLISVYPLIANSVGIYEMKMLAGNARDKKALEAACKATEALEN 904
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 958 IRTVAISLTREDVWEIYHAQDAOGRTSLSVLRSSSLYASSQALVEFCVALGFWYGGTL 1017
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 905 IRTVAISLTQERKESMTYKLEHPEYRNSVQMAHIIYGITFSISQAFMYSVAGCFRGA 964
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1018 L--GHNEY-DIFRFVCESELFGAQSAGIVFSAFPMGRKAKNAAEFRRLPDKRPQIDN 1074
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

```

```

Db 965 IVNHMFRFDV---ILVFSALVFGAVALGHAASSPAPDYAKAKLSAHLFSLFERQPLIDS 1021
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1075 WSEEGKLEFVEGEIERNFVRVPTREPOVFLGLUTVPGQYVALVPGSGGKSTTI 1134
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1022 YSGGLMPDKFEGSVTNEVFNTPRANMPVLQGLSLEFKKGGTLLVSSGGCKSTIV 1081
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1135 ALERFYDALIGSLVDGDKISKLINSYRSFLSVSOEPTLYOGTITKENLIGVEDDV 1194
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1082 QLERFDPDMAGVLLDQGEAKKRLINQMLRQGLIVGSEPTLPFCISAEINAYDGNRNV 1141
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1195 PEEFLIACADANITVDYDMSLPEGFNTVYSGKGMGSGQKQRAVATALLRDKILLD 1254
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1142 SODEIVRAKAAHNPETLPQRYKRVGDKQLSGGQKQRLAIRALLRQPRILLD 1201
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1255 EATSLDSESEKVVQALDAARGTITVAVHRTSTIOKARIVYVFDQKIVESGTHSEL 1314
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1202 EATSLDSESEKVVQALDAARGTITVAVHRTSTIOKARIVYVFDQKIVESGTHSEL 1261
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 1315 VQKGRYYELVNLOS 1329
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1262 LAQGIYFSWVNIQA 1276
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

RESULT 14
DMSIA
multidrug resistance protein 1a - mouse (fragment)
N:Alternate names: P-glycoprotein 1a
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 19-Jan-2001
C:Accession: A34175
R:Hsu, S.I.H.; Lothstein, L.; Horwitz, S.B.
J. Biol. Chem. 264, 12053-12062, 1989
A:Title: Differential overexpression of three mdr gene family members in multidrug-re
A:Reference number: A34175; MUID:89308614; PMID:2473069
A:Accession: A34175
A:Molecule type: mRNA
A:Residues: 1-1104 <RSD>
A:Cross-references: GB:24417; GB:J04839; NID:g200329; PIDN:AAA03243.1; PID:g200330
C:Comment: This is an integral membrane protein overproduced in multidrug-resistant c
tructurally and functionally unrelated lipophilic antitumor drugs.
C:Superfamily: multidrug resistance protein; ATP-binding cassette homology
C:Keywords: ATP; duplication; glycoprotein; nucleotide binding; P-loop; Transmembrane
F:234-458/Domain: ATP-binding cassette homology <ABC1>
F:251-358/Region: nucleotide-binding motif A (P-loop)
F:375-379/Region: nucleotide-binding motif B
F:877-1073/Domain: ATP-binding cassette homology <ABC2>
F:884-901/Region: nucleotide-binding motif A (P-loop)
F:1020-1024/Region: nucleotide-binding motif B
F:257/Binding site: ATP (Lys) #status predicted
F:900/Binding site: ATP (Lys) #status predicted

Query Match 34.7%; Score 2335.5; DB 1; Length 1104;
Best Local Similarity 43.9%; Pred. No. 2.6e-133;
Matches 492; Conservative 203; Mismatches 403; Indels 23; Gaps 5;

```

```

QY 215 DTNLIODGISEKVGITLTLATFYATIAVYKTKLALICSSITVALVLTMGSSQPII 274
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 2 DVSKINIEIGKIKGFQPMATFFGGFIIETFRGKLLVILALISPVGLSAGIMAKLS 61
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 275 KYSKSLDSYAGGTVAEYVSSIRNATAFGTODKIAQOYVHLDAEKWTKNOIVGCF 334
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 62 SEFKELHAYAKAGVAEEVLAARTVYAFAGQKKELERYNNMLEAKRLGKKAITANI 121
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 335 MIGAMFGMYNNGYGFMGSRFLVDGADVGDITLVAMTLIGSFSIGNVSPNQAFTN 394
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 122 SMGAAPFLIYALAFYGTSLVSKYSIGOVLYVFFSVLIGAFSVGQASPNTEARAN 181
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 395 AVAAAKIIFGTIDRSPIDPSYSEKGLDHFEGHIELRNVAHIYPSREVTVMEDVSLM 454
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 182 ARGAAYEYFKIIDNPKSIDSPSKSGHKPDNIQGNLEFNHPSVSKREVOYDIKGLNKKY 241
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 455 PAKTTLVPGSGSKSTIVGVLEVFYMPVKGTYLLDGHDIKDLNLRMQOISLVSOEP 514
   ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

```


Wed Apr 2 09:14:10 2003

us-09-758-828-2.rpr

Page 15

QY	889	TIITSTLCAATIIALAIQIMKTALVCISVPIYLAACGYRFFYMAJOPQSRKLADEGA	948
Db	878	MAVMSFNITIAALLIAEFESWKSLITFTTPEPPLASGAVQTKMLTGRASODKQALERAG	937
QY	949	NFACCATSIRTVASITTRERDVMEIYHAOLDAGRTSLITVLRSSLIYASSQALVFECVA	1008
Db	938	QITSEALSNIRTVAGIGVEGRFKAEEVLEQTSYKAVAKANIYGLCAFEQIAPFLANS	997
QY	1009	LGFEYGGTLLGHHEYDIFRFEVCFSEILCQASAGVSEFADMKAKANAAEFRLDR	1066
Db	998	AANYRGYLLAIYBGLGFSEHVRVYSSVLSATRVGTFESTYTSYAKAKISARFPOILDR	1057
QY	1069	KPOIDNMSEGEKLETVGEIEFRNHFYRPTREPOYVLRGLDTLVKCGQYALVGPSCC	1128
Db	1058	KPPIINYSEAGEKMDNQGRIDPDCFKTFYPSRPIQVNLGSLSVYNGQTLAPFGSSC	1117
QY	1129	GKSTTLLERFDALINAGSLVDGKGISKLTNINSTRFSLSIVSOEPTLYOGTIKENILG	1188
Db	1118	GKSTIQLLERFYDPDQCTVMIDGHSKVMYIQLNSNIGVISOEPVLFDCSIIMNIRYG	1177
QY	1189	IVEDDVPEEFLINACKRANIYDFIMSLPGEFNTVVGSKGMLSGOKORVIAIRALLDP	1248
Db	1178	DNTEIISVEBAIAAAQAOQIHDHVMSLPEKIEFTNNGIQOSSQJLSREKORIAIRAIYVDP	1237
QY	1249	KILLIDEATSALDSESEKVOAALLDAARGRTTIAVAHRLSITQRADYIYFEOCKIYES	1308
Db	1238	KILLIDEATSALDSESEKVOQTALDKARAGRCIYIAHRLSTIQNSDIIVAVSGCVIEK	1297
QY	1309	GTHSEIYOKKGRYIYEV 1325	
Db	1298	THETKILMAOKGATYIKLV 1314	

```
Search completed: April 1, 2003, 16:03:23
Job time : 65 secs
```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:51:33 ; Search time 30 Seconds

(without alignments)
1844.314 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLTNPISPTAMREPAE.....YOKGRYELVNLQSLGKH 1334

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3068.5	45.7	1362	1	PMO1_SCHPO
2	2483	37.1	1276	1	MDR1_CRIGR
3	2489	37.0	1280	1	MDR1_HUMAN
4	2471.5	36.8	1276	1	MDR3_MOUSE
5	2453.5	36.5	1276	1	MDR1_MOUSE
6	2413	35.9	1276	1	MDR2_CRIGR
7	2392.5	35.6	1277	1	MDR1_RAT
8	2367	35.5	1276	1	MDR2_MOUSE
9	2379.5	35.4	1279	1	MDR3_HUMAN
10	2357	35.1	1278	1	MDR2_RAT
11	2357	35.1	1281	1	MDR3_CRIGR
12	2328	34.6	1321	1	AB11_HUMAN
13	2247	33.4	1321	1	AB11_MOUSE
14	2235	33.3	1321	1	AB11_RABIT
15	2230.5	33.2	1321	1	AB11_RAT
16	1985.5	29.5	1321	1	MDR1_CAEL
17	1971	29.3	1302	1	MDR4_MOUSE
18	1886.5	28.1	1254	1	MDR3_CAEL
19	1877.5	27.9	1302	1	MDR5_MOUSE
20	1745	26.0	1280	1	MDR1_LETEN
21	1432.5	21.3	1419	1	MDR1_PLAIF
22	1229.5	18.3	1290	1	STEE6_YEAST
23	956	14.2	1336	1	MAM1_SCHPO
24	927	13.8	715	1	AB10_MOUSE
25	890	13.2	738	1	AB10_HUMAN
26	842	12.5	735	1	AB10_HUMAN
27	830	12.3	1323	1	HST6_CANAL
28	820.5	12.2	762	1	AB10_MOUSE
29	817.5	12.2	762	1	AB10_MOUSE
30	814.5	12.1	766	1	AB10_MOUSE
31	749.5	11.2	582	1	MSBA_ECOLI
32	746.5	11.1	1325	1	MRP4_HUMAN
33	743	11.1	1502	1	MRP6_RAT

34	727	10.8	685	1	MDL1_CANAL	P97998 candida alb
35	724	10.8	1522	1	MRP3_RAT	O88563 rattus norv
36	707.5	10.5	1743	1	TAGC_DICDI	O23868 dictyostel
37	706.5	10.5	1503	1	MRP6_HUMAN	O95255 homo sapien
38	698	10.4	725	1	TAP1_RAT	P26370 rattus norv
39	693	10.3	607	1	HEPA_ANASP	P22638 anabena sp
40	692	10.3	1527	1	MRP3_HUMAN	O15438 homo sapien
41	689.5	10.3	724	1	TAP1_MOUSE	P21958 mus musculu
42	687	10.2	584	1	LMRA_LACTA	O9chl8 lactococcus
43	684.5	10.2	1515	1	YCFL_YEAST	P39109 saccharomyc
44	679	10.1	1436	1	MRP5_MOUSE	O9715 mus musculu
45	678	10.1	1592	1	YHDS_YEAST	P38735 saccharomyc

ALIGNMENTS

RESULT 1
ID PMO1_SCHPO STANDARD; PRT; 1362 AA.
AC P36619; 074513;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-DEC-2002 (Rel. 41, Last annotation update)
DE Leptomycin B resistance protein pmol.
GN PMO1 OR SPCC663.03.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92244046; PubMed=1349418;
RA Nishi K., Yoshida M., Nishimura M., Nishiyama M.,
RA Horiouchi S., Bepu T.;
RT "A leptomycin B resistance gene of Schizosaccharomycetes pombe encodes
RT a protein similar to the mammalian P-glycoproteins";
RL Mol. Microbiol. 6:761-769(1992).
[2]
RP SEQUENCE FROM N.A.
RX SFEAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Collins K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Gellies S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovich E.,
RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,
RA Skellton J., Simmonds M., Squares R., Squares S., Sharp S.,
RA Taylor K., Taylor R.G., Riley A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Roben J., Geyrhofer B.,
RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzjak K., Langer I., Beck A., Leirich E., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Laulau V., Motier S.,
RA Gallbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey A., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Sipkavski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe";
RL Nature 415:871-880(2002).
CC -I- FUNCTION: MAY BE A TRANSMEMBRANE TRANSPORTER OF THE MATING
CC FACTOR, NAMELY P-FACTOR OR M-FACTOR. CONFERS RESISTANCE TO
CC LEPTOMYCIN B AND TO SEVERAL OTHER ANTIFUNGAL DRUGS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC EMBL; D10695; BAA01537.1; -

DR EMBL; AL031307; CAA20363.1; -

DR PIR; S20548; S20548.

DR InterPro; IPR003593; AAA_ATPase.

DR InterPro; IPR003439; ABC_transport.

DR InterPro; IPR001140; ABC_transport.

DR Pfam; PF00005; ABC_tran; 2.

DR Pfam; PF00664; ABC_membrane; 2.

DR ProDom; PD000006; ABC_transport; 2.

DR SMART; SM00382; AAA; 2.

DR SMART; PS00211; ABC_TRANSPORTER; 2.

KW Transmembrane; ATP-binding; Glycoprotein; Antibiotic resistance.

FT DOMAIN 1 91 POTENTIAL.

FT TRANSMEM 92 115 POTENTIAL.

FT TRANSMEM 138 162 POTENTIAL.

FT TRANSMEM 220 237 POTENTIAL.

FT TRANSMEM 244 264 POTENTIAL.

FT TRANSMEM 320 346 POTENTIAL.

FT TRANSMEM 354 374 POTENTIAL.

FT DOMAIN 375 788 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 789 809 POTENTIAL.

FT TRANSMEM 835 859 POTENTIAL.

FT TRANSMEM 916 935 POTENTIAL.

FT TRANSMEM 940 957 POTENTIAL.

FT TRANSMEM 1022 1040 POTENTIAL.

FT TRANSMEM 1054 1072 POTENTIAL.

FT DOMAIN 1073 1362 CYTOPLASMIC (POTENTIAL).

FT NP_BIND 455 462 ATP (POTENTIAL).

FT NP_BIND 1154 1161 ATP (POTENTIAL).

FT CONFLICT 1334 1335 1A -> TC (IN REF. 1).

FT SEQUENCE 1362 AA; 149631 MW; 2ADF8ABE7B3ACEB8 CRC64;

Query Match 45.7% Score 3068.5; DB 1; Length 1362;

Best Local Similarity 47.3% Pred. No. 2.9e-173;

Matches 656; Conservative 233; Mismatches 403; Indels 95; Gaps 18;

QY 14 AMKEPASTTTEBOASTPRADEKKILSLASAPSTTAPADKEHPKSSSSNNNAVSYNE 73

DB 2 SLHSKSTSTVND--NEHSLD---LSIKSLPSNEKNSTKSE--NEASESHVVDYAK 52

QY 74 VDALLAHLEDEKQVLTQLEETKVNISFF-GIMRYATKMDILIMVISTICAIAA----- 127

DB 53 -PFEQYTPDE-DEELIKYKINDTPAKISGPRLISYADKWDIMLOLAGITTGAGLGNP 110

QY 128 -----ASTQRIIMLYQISYDEFDLKNVLYPYVIGIGEFVY 166

DB 111 LMSLYSGQLAQTDLASGKASSFOH-----TVDF-----CLYPTIATIGVFGS 157

QY 167 YVSTVGFIITGEHATQKIREYLESTLRONIGYFDKLAGEVTTTADTNLIDGISE 226

DB 158 YITVWFIIAGERIARIRINQDYLHAILSONIGYFDRLGAGEITRTITTDNFIQDGLGR 217

QY 227 VGLITLALMTFTAFIIVYKWKALLCSSTVALVLTMGSGSOFIIRYKSKSLDSYGA 286

DB 218 VGLVPEFAITPFGVSGVIAIRHKKFTLLISMPAICGIGLGPFTTKTKQIVVAE 277

QY 287 GGVAAEVISTIRNATAFGTODKLAKQYEVHIDEAKMKTKNQIVGFMIGAFGLMYSN 346

DB 278 SSTVEEVESNIRNAPAFGTODLAKLYNKYILTAQRFGINAKIAGLAVGMWFFVAYGV 337

QY 347 YGIGFMGMSRFLVDAVDGDIITVLMATILISFSLGNTSPNAQATNTNVAATAKIFGTI 406

DB 338 YGLAFWEGGRLLAHGDLVDVSKLIGCFPAVLIASYSLANISPKMQSVSCASAAKIFDTI 397

QY 407 DROSPIDPYNECKTLDHEFGHIELRNVAHIYSREPVYMEVSLSPACKTTALVGS 466

DB 398 DRVSPINAFPTGVVWDIGIELEKNIIRVYTPREVLVDNFSLYVCPGKITALVGS 457

QY 467 GSGKSTVGVVERFPVWVREGTVLLDGDHIDIMLRNRQOISVSGPVYFGTTIKNR 526

DB 458 GSGKSTTIGVVERFYDPTIGGVFLDGDKDLTVNASTIRNOSLVQDPVYFATVFEINT 517

QY 527 HGLIGTYENESDQVRELEENAKMANADFTTALPEGETVNGRGFLISGQKORIA 586

DB 518 YGLPDTIKGLSKBELERRRYDAAKLANAYDFTMTPEQSTVNGRGFLMSGQKORIA 577

QY 587 IAAVYSDPKILLIDETSALDTRKSEGVQALERAEGRTTVIAHRLSTITAHNIYV 646

DB 578 IAAVYSDPKILLIDETSALDTRKSEGVQALERAEGRTTVIAHRLSTIRNADNIYV 637

QY 647 LVNGKIAEOGTHDIDVDRGAYRKLVAQNI-----NEQEAALDEADAE-----D 653

DB 638 VNAKTIYEGSHNELLDLNGAVARLVAQKLSGEGKQDEVEELEDADPEIPDTSFGDD 697

QY 694 LTNADIAKIRT-----ASSASDLGKPTTIDRTGTHKSVSAILSKRP-----E 729

DB 698 DEQNDMASTLEAPMASHNTDITLNNKLNKENDVNFEDKTIQHVASEIVNLPADYGEIN 757

QY 740 TTPKYS-----LMTLKFVASFNRP--EIPYMLIGVSVYLAGGQPTQAVLYA 786

DB 758 EEPKRSKSKRNHNEINSTALWFHSFVETMIEICLLIGILASMICGAAYVQAAYFA 817

QY 787 KAISTLSLEPSQSKLRHODFWSLMEFVVGIIOTITQSTNGCAAFVASCRLIRARSTRA 846

DB 818 RFLNITF--DLSTDPDLKAVNFAVYTWLITAIYAFRAVALSNFAMTYAMAVARLRYHL 875

QY 847 FRTILNODIAFDEKENSIGALTSFLETGKHLSGVSGVLTITLMTSTIGAIIITALA 906

DB 876 FRTILNODVFEPRSENVGAIITTSKTQSLGSLGSPITLGFPPQILNITSVTLISLA 935

QY 907 IGMKALVCIISVVPVLIAGCFYFMYLAQFQSKLAYESANFACEATSSIRTVASLNR 966

DB 936 TGMKLGIVLTSTPVITITAGYRVRALDYOETLSAAYESAFAFACESAISRTVASLNR 995

QY 967 ERDVWEIYNAQLOAQRSTLSIVRSLSLYASSQALVFCVALGFWYCGILGHHEDYIF 1026

DB 996 EENVFARVCDSLKPKRESIASLKSGLFSSAAGVTFPLNALTFFVYGSTLMKRGENTYV 1055

QY 1027 RFEVCFSEILFGAOSAGTYEFPADMGKAKNAAEERRLFDKRPQJODNNSEGEKLEYE 1086

DB 1056 QFYTCFLAIVFGIQAGQGFPGYADVTAKKAAAGELIKYLSSEKPKIDTMBECKVYESIQ 1115

QY 1087 -GEIERNVHFRTYPRPEQVLRGLDLVYKPGQYVALVGPBGGKSTTIALERFYDAIA 1145

DB 1116 SAAIEFRQVEFSYPRRHRIKVLRLGLMLTVKPGQVAFVSGSGKSTTIGLIERFYDCDN 1175

QY 1146 GSTLVGDKDISKININSYSFSLVSQEPILVOGTIKENILIGYVDVDEPEELIACAD 1205

DB 1176 GAVLVGVNVRDNIYDKRQALVASQEPILVOGTIRENVIYLG-ASKDVEEEMTEACK 1234

QY 1206 ANIYDFIMSLPEGFNTVYVSGKGLSGQKORVARIARALRDLRDKILLDEATSAIDSESE 1265

DB 1235 ANIHETIIGLPKNYNTLCOGSGSLSGQKORAIARALIRNPKIILLDEATSAIDSSHE 1294

QY 1266 KVVQALDAANAGRTTIAVAHRLSTIQADVIYFDOGKIVESGTHSELVQKGRYELV 1325

DB 1295 KYVQVLEALMAASGRTVVAIAHRLSSIQDADCIYFVDGYIAEAGTHAEIVYKQGRYELV 1354

QY 1326 NQOSLGR 1332

DB 1355 VEGGLNK 1361

RESULT 2

MORL_CRIGR STANDARD; PRT; 1276 AA.

ID MORL_CRIGR

DB 846 ISLIIGWLTLLATLAVPIIATAGVEMKMLSGOALDKKKEGSKITAELENTRTV 905
QY 963 SLTREDDVMEIYHAQDAQRTSLISVLRSLLYASSQALVEFCVAGFWYGGTLGHHB 1022
DB 906 SLTRBQKFNEMVAQSLQITRYRNALKAHVFGITFSTQAMWFSYAACFRGAVLYAREL 965
QY 1023 YQIFRFVCFSELPCQASAGVYFSPAPMGKAKNAAEFRRLFDKRPQIDMSEGEYL 1082
DB 966 MTFENVLVFSAIVFGAMVGVSSFPADYAKAKVSASHIIMEIEVPSIDSTSTGSLRP 1025
QY 1083 ETVEGEIEFRNVRFRYPRPEQPVLRGLDUTYKPGQYVALVYPSGCKSTTIALBFRD 1142
DB 1026 NLEBGVNRKNEVYENPTPRDIPVLOGLMEYKKGOTIALVSSGCKSTVVOILLRFD 1085
QY 1143 AAGSLVDKDKISKLINSYRSFSLVSOEPTLYOGTKEKILGIVEDDPEEFLIRA 1202
DB 1086 PMAGFVFLDGKEVQNLQVOMLRHLGIVSOEPIFLDCSIAENIAYDGNRVSQDEIERA 1145
QY 1203 CADANIYDFINSIPGFRNIVSGSGMLSGGOKORAVARALLRDKIILLDEATSDLS 1262
DB 1146 AEAANIHOPIESLPDKYTRVGDKGTOISGOKORAIARALVROPHIILLDEATSDLT 1205
QY 1263 ESEKVVQALDAARGTITIAVAHRLSTOKADVIYFPOGKIVESGTSSELYOKKGRY 1322
DB 1206 ESEKVVQALDAARRETCITVIAHRLSTIQNDLIVYIONGKYEKGTQOQLAQNGITF 1265
QY 1323 ELVNTSLQK 1332
DB 1266 SMVSQVQAK 1275

RESULT 3
MDRL_HUMAN
ID MDRL_HUMAN STANDARD; PRT; 1280 AA.
AC P08183; Q12755; Q14812;
DT 01-MUG-1988 (Rel. 08. Created)
DT 01-NOV-1997 (Rel. 35. Last sequence update)
DT 15-JUN-2002 (Rel. 41. Last annotation update)
DE Multidrug resistance protein 1 (P-glycoprotein 1) (CD243 antigen).
GN ABCB1 OR PGY1 OR MDRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87028230; PubMed=2876781;
RA Chen C.-J., Chin J.E., Ueda K., Clark D.P., Pastan I., Gottesman M.M.,
RA Robinson I.B.;
RT "Internal duplication and homology with bacterial transport proteins
in the mdrl (P-glycoprotein) gene from multidrug-resistant human
cells.";
RL Cell 47:381-389(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90094448; PubMed=1967175;
RA Chen C.-J., Clark D.P., Ueda K., Pastan I., Gottesman M.M.,
RA Robinson I.B.;
RT "Genomic organization of the human multidrug resistance (MDRL) gene
and origin of P-glycoproteins.";
RL J. Biol. Chem. 265:506-514(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97190336; PubMed=9038218;
RA Chen G., Duran G.E., Steger K.A., Lacayo N.J., Jaffrezou J.P.,
RA Dumontet C., Sikik B.I.;
RT "Multidrug-resistant human sarcoma cells with a mutant P-glycoprotein,
altered phenotype, and resistance to cyclosporins.";
RL J. Biol. Chem. 272:5974-5982(1997).
RN [4]
RP SEQUENCE OF 1-234 FROM N.A.
RA Smith A., Beck C., Gibson A.;

RL Submitted (Aug-1997) to the EMBL/Genbank/DBJ databases.
RN [5]
RP SEQUENCE OF 178-215 AND 800-856 FROM N.A.
RX MEDLINE=90290529; PubMed=1972623;
RA Gekeler V., Weger S., Probst H.;
RT "mdrl/P-glycoprotein gene segments analyzed from various human
leukemic cell lines exhibiting different multidrug resistance
profiles.";
RL Biochem. Res. Commun. 169:796-802(1990).
RN [6]
RP SEQUENCE OF 1-23 FROM N.A.
RA Kioka N., Tsubota J., Kakeni Y., Komano T., Gottesman M.M.,
RA Pastan I., Ueda K.;
RL Submitted (Jul-1991) to the EMBL/Genbank/DBJ databases.
RN [7]
RP VARIANTS SER-893 AND THR-893.
RX MEDLINE=21686803; PubMed=11829140;
RA Saito S., Iida A., Sekine A., Mura Y., Ogawa C., Kawachi S.,
RA Higuchi S., Nakamura Y.;
RT "Three hundred twenty-six genetic variations in genes encoding nine
members of ATP-binding cassette, subfamily B (ABCB/MDR/TRP), in the
Japanese population.";
RL J. Hum. Genet. 47:38-50(2002).
CC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC - SUBCELLULAR LOCATION: Integral membrane protein.
CC - TISSUE SPECIFICITY: EXPRESSED IN LIVER, KIDNEY, SMALL INTESTINE
CC AND BRAIN.
CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC - DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chromancer/Genes/PGY1ID105.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
CC or send an email to license@isb-sdb.ch).
CC
DR EMBL, M14758; AAA59575.1; -;
DR EMBL, M29447; AAA59576.1; -;
DR EMBL, M29447; AAA59576.1; JOINED.
DR EMBL, M29445; AAA59576.1; JOINED.
DR EMBL, M29426; AAA59576.1; JOINED.
DR EMBL, M29427; AAA59576.1; JOINED.
DR EMBL, M29428; AAA59576.1; JOINED.
DR EMBL, M29429; AAA59576.1; JOINED.
DR EMBL, M29430; AAA59576.1; JOINED.
DR EMBL, M29431; AAA59576.1; JOINED.
DR EMBL, M29432; AAA59576.1; JOINED.
DR EMBL, M29433; AAA59576.1; JOINED.
DR EMBL, M29434; AAA59576.1; JOINED.
DR EMBL, M29435; AAA59576.1; JOINED.
DR EMBL, M29436; AAA59576.1; JOINED.
DR EMBL, M29437; AAA59576.1; JOINED.
DR EMBL, M29438; AAA59576.1; JOINED.
DR EMBL, M29439; AAA59576.1; JOINED.
DR EMBL, M29440; AAA59576.1; JOINED.
DR EMBL, M29441; AAA59576.1; JOINED.
DR EMBL, M29442; AAA59576.1; JOINED.
DR EMBL, M29443; AAA59576.1; JOINED.
DR EMBL, M29444; AAA59576.1; JOINED.
DR EMBL, M29445; AAA59576.1; JOINED.
DR EMBL, M29446; AAA59576.1; JOINED.
DR EMBL, AF016535; AAB69423.1; -;
DR EMBL, AC002457; AAC82531.1; -;
DR EMBL, M37724; AAA88047.1; -;
DR EMBL, M37725; AAA88048.1; -;
DR EMBL, X58723; CAA41558.1; -;
DR PIR, A25059; DVH01.
DR PIR, A34914; A34914.
DR Genew, HGNC:40; ABCB1.

DR MIM: 171050; AAA_Atpase.
 DR InterPro: IPR003593; AAA_Atpase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABCtransport.
 DR Pfam: PF00005; ABC_tran. 2.
 DR Pfam: PF00664; ABC_membrane. 2.
 DR ProDom: PD000006; ABC_transport. 2.
 DR SMART: SM00382; AAA. 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family; Polymorphism.
 FT DOMAIN 1 51
 FT TRANSMEM 52 72
 FT TRANSMEM 120 140
 FT TRANSMEM 189 209
 FT TRANSMEM 216 236
 FT TRANSMEM 297 317
 FT TRANSMEM 326 346
 FT TRANSMEM 347 710
 FT TRANSMEM 711 731
 FT TRANSMEM 757 777
 FT TRANSMEM 833 853
 FT TRANSMEM 854 874
 FT TRANSMEM 937 957
 FT TRANSMEM 974 994
 FT DOMAIN 995 1280
 FT NP_BIND 427 434
 FT NP_BIND 1070 1077
 FT REPEAT 1 637
 FT REPEAT 638 1280
 FT CARBOHYD 91 91
 FT CARBOHYD 94 94
 FT CARBOHYD 99 99
 FT VARIANT 893 893
 FT VARIANT 893 893
 FT VARIANT 893 893
 FT CONFLICT 23 23
 FT CONFLICT 185 185
 FT CONFLICT 336 336
 FT CONFLICT 412 412
 FT CONFLICT 438 438
 FT SEQUENCE 1280 AA: 141462 MW: 481675.143675 CRC64;
 Query Match 37.0%; Score 2489; DB 1; Length 1280;
 Best Local Similarity 42.1%; Pred. No. 4,5e-139; Mismatches 448; Indels 60; Gaps 9;
 Matches 555; Conservative 229; Mismatches 448; Indels 60; Gaps 9;
 QY 94 EELKVNISFGLMRVATKMDILIMVISTICAIAAASTFORIMLY----- 137
 DB 28 KERKPTVVSFMRYSNMDLKLVMVVTLLAIIHGAGLPLMLLVGEMTDIFANAGMLED 87
 QY 138 -----QISYDFEY---DELTKNVLYEVLGIGEFYTVVSYVGFYITGHAOKI 184
 DB 88 LMSNITNSDINTGFFMNEEDMTRAYVYSGIAGVLAAYTIQVSFWCLANGROTHKI 147
 QY 185 REYLLSILRNQNGYFDKLCAGEVTTITADTLMIDIGISEKGLTTLATLFTATFATIIA 244
 DB 148 RKDFPFAHMQELGWDVHYGELNRLTLDVSKINIGIDKIGMFQSNATPTGTIGV 207
 QY 245 YKRYKALICSTVALVLTMGSGSOFIYKSKLSIDSYAGAGTVAEYVSIIRNATF 304
 DB 208 FTRGMKLTITVLAISPLGLSAVWMAKLISSFTDKELLAVAKAGAVEVLAIRVIATF 267
 QY 305 GTQDLAKQVEVHLDKEKGTKNQIYMGPMIGAMGLMSNGLGWMSRPLVDCAVD 364
 DB 268 GGGKKELERNNKLEKKRIGIKAITANTISGAFLIYASVALAWYGTITVLSGEYS 327
 QY 365 VQDILVYMAIILGFSIGVSNPAQFTNAVAAAIAKIFGTIPIROSLDIPYSNGKTLHD 424
 DB 328 IGVVLYVFEVVLGAFSGVSGASPIEFANARGAAYEIFIIDKPSIDSYSGHAPDN 387
 QY 425 FEQHIELRNKHIYPSRPEVTYMEDVSLSMAPAGKTALVGPSSGKSTVVGLEVFYMPV 484

DB 388 IKGNLEFRNVHFSYSRKEKILKGLMINKVSGGQVVALVNSGGKSTVQOLMRLDPT 447
 QY 485 RGVLLDGHIDKIDNLNMLRQQLSVASOEPLFETTYIKNRHLLIGTKYENSEEDVRE 544
 DB 448 EGVSVGDDQRTNVTNFFLREIIGVSOEPLVETIAENTRYG-----RNVMTDE--- 499
 QY 545 LIENAAKMANAHDTTALPEGETNWGQGFLLSGQKORAIARAVVSPKILLDEAT 604
 DB 500 -IEAAVEANAYDRIKMLPKHFDTLVBERGQQLSGQKORAIARALVNRNPKILLDEAT 558
 QY 605 SALDTKEGVVQALERPABEGRTTIVAHRLSTIKTAHNVLVNKGIAEGTHDELVD 664
 DB 559 SALDTEEAAYVQALDKARGRRTTIVAHRLSTVRNADVLAGFDGVYVERGNDELK 618
 QY 665 GGAARKIVEORINEQKADALDEADDLTNADIAIKTASSASPDLDKPTTIDTGT 724
 DB 619 KGIEFKLVMTQAGNEVE--LENA-----ADESKSEIDALEMSSNDSSILRRST 668
 QY 725 HKVSSAAILSKRP-----PETPKYSIMTLKFKVASFNRPEIPLYMIGLVFVLGGG 777
 DB 669 RRSVRSQADRKISTKEALDESTIPVSPFRIMK-----INTEPPTVVGVCALINGCL 724
 QY 778 QPTQAVLYAAISTLSLPESQYSKLRHDADEFWSLMFEVVGIIQETQSTNGAFAVCSER 837
 DB 725 QPAPAIIFSRLIGVETRIDDETRK-RQNSMLFSLFALAGIISFTIFELQGFTEGKAGEI 783
 QY 838 LIRARSTARTTILROIAFPDKENSTGALTSLSLSTETKLSGVSGVTTGTLMTSTTL 897
 DB 784 LTKRLRYMFRSMRLROVSWEDDPKNTTGALTYTLDAADQVKCAISRLAVIOTNANTL 843
 QY 898 GAATIIIALGKMLALCISVVPVLLACGFYRFYMLAOFQSRSLKAVEGSAFACETASS 957
 DB 844 CTGIIISFTIGWQTLTLLLAIVPIIALAGVEMMLSGQALKDKELEGAKIATELEN 903
 QY 958 IRTVASTFRDWEIYHAQDAQRTSLISVRSLSLASSQALVEFCVALGFWYGTTL 1017
 DB 904 FRIVVSLTQDKFEHMTAOSLOVYRNSLRKAHIFTFSFTQMTMFTSVAGCFRCAYL 963
 QY 1018 LGHHEVYIFRPVFESEILGAGASGVTSFAPDMGAKNAAEFRLPDRKPIDMWSE 1077
 DB 964 VAHKLMSFEDVLVFSVVGMAVAGVSPAPYAKAKISAAIIMIEKTPIDIDYST 1023
 QY 1078 EGEKLTVEGEIEFRNHFRTPRPOVLRGLDLYKPGQVVALVGPSSGKSTTALL 1137
 DB 1024 EGLMPTNLEGNVTGEVFNTPRPDLPVLQGLSLEKKGOTLAVSSGCGKSTVQL 1083
 QY 1138 ERFYDAIAGSLVDGKDISKLINISYRSLVSQEPFLYOGITKENILLGIYEDVPEE 1197
 DB 1084 ERFYDPLAGKVLIDGKEIKRLNVQMLAHLCIYVSEPIFLFDCSIAMENIYAGDNGRVVSOE 1143
 QY 1198 FLTKACDANIDYRIMSLPEGFNTVVGSKGMLSGQKORAIARALLRPKILLDEAT 1257
 DB 1144 EYVAAKEANINHATESLIPKYSITKVGDKGQLSGGKORAIARALVROPHILLDEAT 1203
 QY 1258 SALDSESEKVVQALDAARGRTTIVAHRLSTQKQADVIVVPOGIVSGTSELVOK 1317
 DB 1204 SALDTESEKVVQALDARREGRTCIYAHRLSTIQNADLIVFONGKVEKHTHQLLAQ 1263
 QY 1318 KGRYVELVNTQS 1329
 DB 1264 KGIFYSMVSVOA 1275
 RESULT 4
 MDR3_MOUSE
 ID MDR3_MOUSE STANDARD; PRT; 1276 AA.
 AC P21447;
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Multidrug resistance protein 3 (P-glycoprotein 3) (MDR1A).
 GN ABCB1A OR ABCB4 OR PGY3 OR PGY-3 OR MDR3 OR MDR1A.

Db 783 RLRYVFKSMKRODVSWEDDPKNTTGALTTRLANDAAQVKGATGSLAVIFQNTIANITG 842
 QY 901 IITAIAGMIALVCISVPEVLACGFREFYMLAQFQSRKLAEGSANACEATSSIRT 960
 Db 843 IITSLIYQMOTLLIATVYIITAGVCEKMLSGOALKRKEGSKATLAEIEMFRT 902
 QY 961 VASLTREBDWEIYHQAQDAGRTSLISVRSLLVASSQALVEFCVALGFYWGTLGH 1020
 Db 903 VASLTREBDWEIYHQAQDAGRTSLISVRSLLVASSQALVEFCVALGFYWGTLGH 962
 QY 1021 HEVDIRFVFCSEILFGAASACTVPSFARDMGAKANAAEFRLFRKQIDWMSSEGE 1080
 Db 963 QLTFFENVLLVFAIYGAAGVQSSFPADYAAATVSAASHIIRIETKPEIDISTQGL 1022
 QY 1081 KLETFEIEIFRWRHFRPTPEQPVLRGLDLYVKGQYVALVPSGCGSTTALLERF 1140
 Db 1023 KPMLEBNQVQSGFVFPYPRPSPVQGLSLEKKGOTLAVSSCGSTVQQLERF 1082
 QY 1141 YDAIAGSILVDGDISKLNINSYRSFLSLVSOEPTLYOGTKENILLGIYEDDVEEFLI 1200
 Db 1083 YDMAGSVFLDGKEIKQALQVMLRAQGIYSEBILFDCSIAENIAYGDSRVSYSEIY 1142
 QY 1201 KAKDANIVYFIMSLPGEFTVVGSKGMLSGCKORVATARALLRDKILLDEATASL 1260
 Db 1143 RAKEANIHQFIDSLPRKYTRVGDGTQLSGCKORVATARALLRDKILLDEATASL 1202
 QY 1261 DSESEKVAQALDPAAGRTTIAVHRSLTQKADVIYVFDGKIVSGTSELSVOKKR 1320
 Db 1203 DTSEKVVQALDKARGRCIYIAHRLSTIQNADLIVYLONGKVRKHGHQQLAOKGI 1262
 QY 1321 YVELVNLQSLGK 1332
 Db 1263 YFSNVSVQAGAK 1274

RESULT 5
 MDRL_MOUSE
 ID MDRL_MOUSE STANDARD; PRT; 1276 AA.
 AC P06795;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Multidrug resistance protein 1 (P-glycoprotein 1).
 GN ABCB1 OR ABCB1 OR PGY1 OR PGY1-1 OR MDRL OR MDRLB.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=87028229; PubMed=3768958;
 RA "Gros P., Croop J., Housman D.;
 RT "Mammalian multidrug resistance gene: complete cDNA sequence
 RT indicates strong homology to bacterial transport proteins.";
 RL Cell 47:371-380(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=89367274; PubMed=2570420;
 RA Raymond M., Gros P.;
 RT "Mammalian multidrug-resistance gene: correlation of exon
 RT organization with structural domains and duplication of an ancestral
 RT gene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:6488-6492(1989).
 RN [3]
 RP SEQUENCE OF 1-21 FROM N.A.
 RP MEDLINE=91042535; PubMed=2248681;
 RA Raymond M., Gros P.;
 RT "Cell-specific activity of cis-acting regulatory elements in the
 RT promoter of the mouse multidrug resistance gene mdrl.";
 RL Mol. Cell. Biol. 10:6036-6040(1990).
 CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
 CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
 CC LINKER DOMAIN.
 CC -1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
 CC RELATED BUT DISTINCT CELLULAR GENES.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M14757; AAA79005.1; -;
 DR EMBL: M60348; AAA39513.1; -;
 DR PIR: A33719; DVM51.
 DR MGI: 97568; Abcd1b.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABCtransport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR ProSite: PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family; Phosphorylation.
 FT DOMAIN 1 47
 FT TRANSSEM 48 71 POTENTIAL.
 FT TRANSSEM 119 139 POTENTIAL.
 FT TRANSSEM 199 219 POTENTIAL.
 FT TRANSSEM 296 316 POTENTIAL.
 FT TRANSSEM 327 347 POTENTIAL.
 FT DOMAIN 348 709 POTENTIAL.
 FT TRANSSEM 710 730 POTENTIAL.
 FT TRANSSEM 755 775 POTENTIAL.
 FT TRANSSEM 831 851 POTENTIAL.
 FT TRANSSEM 855 882 POTENTIAL.
 FT TRANSSEM 935 955 POTENTIAL.
 FT TRANSSEM 966 986 POTENTIAL.
 FT DOMAIN 987 1276 POTENTIAL.
 FT NP_BIND 426 433 ATP (BY SIMILARITY).
 FT NP_BIND 1068 1075 ATP (BY SIMILARITY).
 FT REPEAT 1 635
 FT CARBOHYD 636 1276
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1276 AA; 140993 MW; 1804D0F011B0FF4E CRC64;

Query Match 36.5%; Score 2453.5; DB 1; Length 1276;
 Best Local Similarity 40.9%; Pred. No. 5.6e-137;
 Matches 520; Conservative 232; Mismatches 461; Indels 57; Gaps 7;
 QY 90 KTOLEIKVNISFGLMRYATKMDLIVISTICIAAASFFQRLM----- 136
 Db 23 KKEKKEKPAVGVGMFRYADMIDKMLITGLAIHGTLLPLMLVFGMNTDSFTAE 82
 QY 137 -----XOISYDEFYDELTKNLYFYVYLGIEFYVYVYVSTGVITYGEHA 180
 Db 83 ASILPSTINGSPNSTLIISNSLEEMAVIAYYYTGAGVLIYAVIQVSLMCLAAQRO 142
 QY 181 TOKIREYLESILRONIGYFPKLAGEVYTRITADTNLIDGISKVGLTALATFPTA 240
 Db 143 IHKIRQKFFHAIMQELIMFVHDVHGEINTRLTDVSKINDIGIKIMFQSTITFLAG 202
 QY 241 FIIVAVKKLALICSSITVALVLTLMGSGSOFIIRYSSKSDYAGGVAAEVVSSIRN 300
 Db 203 FLIGISCKLTVLAVLSPILGLSSALMAKVLFTTKELQAYAKAAVAAEVLAART 262
 QY 301 ATAFTQDLAKQYEVHLDDEAKMGTKNQIYVGFIMGAFGLMYSNYGLGFWMGSRFLVD 360

```

Db 263 VYAFGGQKLEERNNKLEAKNGVKKALTAITASISIGIAVLVAYASTALFWGTSLVLS 322
QY 361 GAVVDGDIIVLVAAILIGSFSLNVPNAOFTNAVAAAARFCTIDROSPLDPYSECK 420
Db 323 NEXSIEGVLVFVSILGTSGIHLAPNIEAFANARGAPEIRKIIDNESIDSFISKGY 382
QY 421 TLDFBCHIELRNVKHYPSRPEVYEDVSLSPACKTTALVPGSGSKSTVGLVERF 480
Db 383 KPDSINGNLEFKVHFVPSRSEVOILKGLNKYSGQVALVAGNSGCKSTVQLMQL 442
QY 481 YMPVREGVLIDGDIDKIDNRMLRQOISLVSQEPVLEGTIYNIRHGLGTKEENSED 540
Db 443 YDPLEGVASTIDGODIKTINRYLREITIGVVSQEPVLEATTIENIRG-----RED 493
QY 541 KYBELIENAKMANADFITALEPEGETYNGORGFLLSGOKORAIARAVSPDKILL 600
Db 494 VTMDELEKANKKANAYDFIMKLPHQDPTLVGEGAGQSLSGOKORAIARALVKNPKILL 553
QY 601 DEATYSLDTRKSEGVQAALERRAEGRTTYIAHRLSTIKTAHNIIVLVNCKIAEGTHDE 660
Db 554 DEATYSLDTRKSEGVQAALERRAEGRTTYIAHRLSTIKTAHNIIVLVNCKIAEGTHDE 613
QY 661 LVDRGAYRVLVEAO--RINQKREAD-----ALEBDAEDLTNADIARKTKTASSSDLDGK 715
Db 614 LMKREKIFYELVMTQTRNGNIEPGNNAYGSDPTDASELTSESKSPLIRRSIYRSHRK 673
QY 716 PTDIDRTGTHKSSAISILSRPETTPKYSLMTLTKFVAFNRPETPYMLIGVSVLAG 775
Db 674 QOQERLSMKEND-----EDVPVVSFWIL-----NLNISEMPYLVAVGLCAVING 720
QY 776 GGQPTAVLYAKAISTLSPESOYSKLRHDADWFLMFFVVGIIQFTOSTNGAARVCS 835
Db 721 CIOPTAVLYAKAISTLSPESOYSKLRHDADWFLMFFVVGIIQFTOSTNGAARVCS 779
QY 836 ERLIRARSAFTILRODIAFEDKENSGLATSLSTETKILSGSVTGLTILMTST 895
Db 780 EILTKRVRVWFVSMKRODISWPDHNSGSTLTRIASASSVKGAMGRALVAVQNA 839
QY 896 TLGAAIILALATGKIALYCVIPVLLACGFYFYMALDFOFSRKLYAGSANFACAT 955
Db 840 NGTGVILSLVGMQLLVLVILPIVIGIIEMLKLSQALKDKOLEISKINTEAI 899
QY 956 SSTRVASTLRWDVEIYHAQIDAGRTSLISLVNSSLVYASSQALVFECVALGFWYGG 1015
Db 900 ENFRIVSLTRQKFTMTYAQSLQVRYRNAMKRAHYGFIFFSTQAMMYSTYACFRGA 959
QY 1016 TLIGHHEYDIFRFVCFSEILFGAOSAGTVFSAPDMGKAKNAAEFRRLFDKPKIDNM 1075
Db 960 YLVAQOLMFFENMLVFSAVFGAMAGNTSSFAPIYAKAKVASHIRIETKPEIDSY 1019
QY 1076 SEEGKLETVEGEIEFRNHFRTYRPEQVLRGLDLTVKQGYVALVPGSGCKSTTIA 1135
Db 1020 STEGKLPITLLEGVKNVNGVFNPTRPNIPLYGLSLFKKQGLTALVSGSGCKSTVQ 1079
QY 1136 LLERFDAIAGSLVNGKDISKINISYFSLVSQEPVLYGCTIKENTILGIVEDDV 1195
Db 1080 LLERFYDPNAGSVFLDGKEIKQLNVOMLRAHLGIVSQEPILFDCSIAENIAYDNSRAVS 1139
QY 1196 EEFILACDANITDFITMSLPEGFNTVSGSGKMLSGGOKORAIARALRDKILLDE 1255
Db 1140 HEEIYAAKEANIHOFIDISDPDKYNTRVGDKQLSGGOKORAIARALVROPHILLDE 1199
QY 1256 ATSSALDSESEKVVQALDAARGRITIAVHRLSTIQKADVIVPQGVESGTSISELY 1315
Db 1200 ATSSALDSESEKVVQALDAARGRITIAVHRLSTIQKADVIVPQGVESGTSISELY 1259
QY 1316 QKGRYELV 1325
Db 1260 AOKGIFSMY 1269

```

```

MDR2_CRIGR
ID MDR2_CRIGR STANDARD; PRT; 1276 AA.
AC P21449;
DT 01-NOV-1991 (Rel. 18, Created)
DE 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 2 (P-glycoprotein 2).
GN PGY2 OR PGP2.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_Taxid=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
  gene family.";
RL DNA Seq. 2:89-101(1991).
RN [2]
RP SEQUENCE OF 622-1276 FROM N.A.
RX MEDLINE=88122132; PubMed=2893225;
RA Endicott J.A., Uraoka P.F., Sarangi F., Gerlach J.H., Deuchars K.L.,
  Ling V.;
RT "Simultaneous expression of two P-glycoprotein genes in
  drug-sensitive Chinese hamster ovary cells.";
RL Mol. Cell. Biol. 7:4075-4081(1987).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
  DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
  CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
  CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60041; AAA68884.1; -
DR EMBL: M17896; AAA37007.1; -
DR PIR: B27126; DVHY2C.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABC_transporter.
DR Pfam: PF00005; ABC_tran. 2.
DR Pfam: PF00664; ABC_membrane. 2.
DR ProDom: PD000006; ABC_membrane. 2.
DR SMART: SM00382; AAA. 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
  Multigene family.
FT DOMAIN 1 51
FT TRANSMEM 52 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 119 139 POTENTIAL.
FT TRANSMEM 188 208 POTENTIAL.
FT TRANSMEM 215 235 POTENTIAL.
FT TRANSMEM 296 316 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT DOMAIN 346 708 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 709 729 POTENTIAL.
FT TRANSMEM 755 775 POTENTIAL.
FT TRANSMEM 831 851 POTENTIAL.
FT TRANSMEM 852 872 POTENTIAL.
FT TRANSMEM 935 955 POTENTIAL.
FT TRANSMEM 972 992 POTENTIAL.
FT DOMAIN 993 1276 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 426 433 ATP (POTENTIAL).

```

RESULT 6

FT NP_BIND 1068 1075 ATP (POTENTIAL).
 FT REPEAT 1 635
 FT REPEAT 636 1276
 SQ SEQUENCE 1276 AA; 141057 MW; 5096B1385628812D CRC64;
 Query Match 35.9%; Score 2413; DB 1; Length 1276;
 Best Local Similarity 40.8%; Pred. No. 1,4e-134;
 Matches 520; Conservative 232; Mismatches 465; Indels 56; Gaps 10;
 86 RYKLTQLEIKYINISFGLMRYATKMDILINVTSTCAIAASTFORIMLY----- 137
 20 RKSKKKEKENENNGVGFGRADMDIKLYMGLTAAVLHGTSLPLMLVGNMTDSF 79
 138 -----QISYDEYDELFRNVLYFYVIGIEFYVTVSVTYTG 177
 80 TRAKTSIMPMTNQSEINNEVVISGLSEDMATVAYYTGIGAGVLVAVIYQVSWCLAA 139
 178 EHAQTQIREYLESILRONIGYFDKLGAGVYTRITADPLIDGJSEKVGILLTALATP 237
 140 GRQINKIRQKFFFAIMNOEIGMPVDHIGELNRLTDVSKINDGIDKIGMEFQSIATP 199
 238 VPAFIATVYKMTALICSTVALVYTMGGSGOETIKYSKLSDSYAGAGTAAEEYISS 297
 200 LAAFYVGFISGKMLTYIILAVSPILIGLSAMAKVLTSPFNKELQATKAGAAVEYLAA 259
 298 IRNATAFGTQDKLAKOYEVHLDAEKWGTNOJYMGFMIGLMSNYGLCFMWSRF 357
 260 IRFVIAFGQNKELERKNKMLEAKNVGIRKAVTANISIGIAYLLVYASVALAFWYGTSL 319
 358 LVYGAADVGLIYVLAAILIGSISLVNSPNAQAFNAVAAAKIFCTIDROSLLDPSN 417
 320 VLSNEYSVGVLLVFSILGTSIGHIAPINEFANARAAAEFIITIDNEPSDFST 379
 418 EGKTLDFEHIERNKHYIPSPREYTVMEDVSLSPACKTALVPSGSGKSTYVGLV 477
 380 QGHKPSVGNLFKKNVHFPSRSKGIILKGLAKVOSQYALVYKSGCGKTYQLL 439
 478 ERFYMPVGVLLDGDHDKDLNMLKQOISLVSOEPVLEFGTITKKNIRHGLIGTKYNE 537
 440 QRLYDPFEGVYSIDGDIRTINRYLREIIGVYSQEPVLEFATTAENIRG-----RENV 494
 538 SEDKVRLEINAKMANAHFITALPGSYTNVQORFLLSGCGKOKIATARAAYSPKI 597
 495 TMBE---IKKAVKAAAYDFIKLPHKFDPLVGERAOISGGOKRIATARAAYSPKI 550
 598 LILDEATSLDTSSEGVAALEERAEGRTIYIAHRLSTIKYTHNIVLVYNGKIAOGT 657
 551 LILDEATSLDTSSEAVVQALDKARGRITTYIAHRLSTIVRNADVLAGFDGVIYQGN 610
 658 HDELVDGAYRRLVEAQ-RINQK---EADALE-DADAEDLTNADIATIKITASASSDL 712
 611 HEELMEKRGYICRLVMYQGTGNEVELGSEADGSDSTIASELTSEERKSPSVKSTORSI 670
 713 DGRPTIDRGTIKSVSALLSKRPPTPKYSLMTLLKAVASNNRREIYMLIGLVSV 772
 671 CGSQDQDERRV---SVKEA-----QDDVDPLVSWFKILK---LNTIEWPLVGVICAV 717
 773 LAGGQPTQAVLAKAISTSLPESQSKLRHADFWSLMEFYVGIOTFQSTNGAFA 832
 718 INCGMQPVEIVSGIIGVTRDDPKTK-QQNCNLFSLFPLVGMICFTYTFQGTGFC 776
 833 VCSERILRRASRTAFRTILRODIAFAFDKENSFGALNFTSTETKHLGVSQVTLGTLIM 892
 777 KAGEILLKRLRYMFKSMLRODISMFDHNSGALTTRILASDAANVKGAMSSLAGITQ 836
 893 TSTTLGAITIALAIGKALAVCSVYPVILACGFYFVMAQOSKSLAYESANFAC 952
 837 NVANLGGIILSVYQQLLLLVAVLPIILSGMEMEKVLSGQALDKLELVSGIAT 896
 953 EATSIPTVASLFRERPMVEIYHQAODAGRTSLISVRSLSLVAASQALVEFCVALGFW 1012
 897 EAIENFTVSLTRQKFFEMKYAOSLOIPYRNALAKAHVIGITFSFTQAMMYFSYAACFR 956

QY 1013 YGGLTLCHEHYDIFRFVCSSELLFGAOSAGTVFSPADMGKAKNAAEFRRLDFRRQI 1072
 DB 957 FAYLVVAHQIMTEENVALVESAVVFGAIAAGNASSFPADAKAVASHTIRIEKIPSI 1016
 QY 1073 DNMSEGEKLETVGCELEFNNVHRFPYTPROPLRBLDLVYRGQVVALVPGSGCKST 1132
 DB 1017 DSVTRKLKPMWEGNKKFMEVFNFPYPRDIPPLQGLSLKVKGQTLALVGGSGCKST 1076
 QY 1133 TIALREFYDALAGSILVDEKDISKLINSTRFSLVSQEPFLYQGTIKENILGIYED 1192
 DB 1077 VQOLLEFPYPMAGTVLDDKEIKQLVQVNRALHIGVSOEPILFDCSINENIAYGNSR 1136
 QY 1193 DVPEEFLKACKDANIYDFIWSLPEGFNTVYSGKMLSGGQKQVARIARALLRDKILL 1252
 DB 1137 VVSODEIERAKAEKANIHFIESLPDKYNTVRGDKGTQLSGGQKORIAARALVROPHILL 1196
 QY 1253 LDENTSLDSESEKVVQALDAARGTTTAVARLSTIQKADVIYFPDQKIVESGTHS 1312
 DB 1197 LDENTSLDSESEKVVQALDAARGRTCIYARLSTIQKADVIYFPDQKIVESGTHS 1256
 QY 1313 ELVQKRGYELV 1325
 DB 1257 QLAQKGIYFSMV 1269
 RESULT 7
 MDR1_RAT
 ID MDR1_RAT STANDARD; PRT; 1277 AA.
 AC P43245;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Multidrug resistance protein 1 (P-glycoprotein 1).
 GN ABCB1 OR P-gly OR MDR1 OR MDR1B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92039081; PubMed=1682220;
 RA Silverman J.A., Raunio H., Gant T.W., Thorgelirsson S.S.;
 RT Cloning and characterization of a member of the rat multidrug
 RT resistance (mdr) gene family.*;
 RL Gene 106:229-236(1991).
 CC - FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
 CC - DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
 CC - SUBCELLULAR LOCATION: Integral membrane protein.
 CC - PTM: SEVERAL PHOSPHORYLATED SERINE RESIDUES ARE PRESENT IN THE
 CC LINKER DOMAIN.
 CC - MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE
 CC RELATED BUT DISTINCT CELLULAR GENES.
 CC - SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M81855; -; NOT ANNOTATED CDS.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABC_transport.
 DR Pfam: PF00064; ABC_tran; 2.
 DR Pfam: PF00064; ABC_membrane; 2.
 DR Prodom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family.

FT	DOMAIN	1	47	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	48	71	POTENTIAL.
FT	TRANSMEM	119	139	POTENTIAL.
FT	TRANSMEM	199	219	POTENTIAL.
FT	TRANSMEM	296	316	POTENTIAL.
FT	TRANSMEM	327	347	POTENTIAL.
FT	DOMAIN	348	709	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	710	730	POTENTIAL.
FT	TRANSMEM	755	775	POTENTIAL.
FT	TRANSMEM	832	852	POTENTIAL.
FT	TRANSMEM	854	874	POTENTIAL.
FT	TRANSMEM	937	957	POTENTIAL.
FT	TRANSMEM	968	988	POTENTIAL.
FT	DOMAIN	989	1277	CYTOPLASMIC (POTENTIAL).
FT	NP_BIND	426	433	ATP (BY SIMILARITY).
FT	NP_BIND	1070	1077	ATP (POTENTIAL).
FT	REPEAT	1	635	
FT	REPEAT	636	1277	
FT	CARBOND	91	91	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOND	96	96	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SEQUENCE	1277 AA;	141386 MW; 8AFDD6D19D2934C1 CRC64;	
 Query Match				
Best Local Similarity		35.6%;	Score 2392.5;	DB 1; Length 1277;
Matches 513; Conservative		40.4%;	Pred. No. 2.3e-133;	
		236;	Mismatches 459;	Indels 61; Gaps 10.
Qy	94	EIKKINISFFGLMRATKMDILINVIISTCAIAASTFORIML-----	136	
Db	26	KKKRPAVGIFGFRVADMLDKLMALGTALIIHGTLPLMLLVGYMTDSPTSPNDPS	85	
Qy	137	-----YQISTDFIDELTKNLYRVYVIGIGEFVTYYVSVCFFITGGEATOK	183	
Db	86	DRAITNOSEINSTHVAWSDTSLIEDMAMAYVYTGTAGAVLIYAAYIOVSLMCACGRHQIK	145	
Qy	184	IREYLESILRONIGSYFDKLAGEWTTRTPADTNLIODISEKVGTLTALATFVYAFIT	243	
Db	146	IRKFHFHIMNDELGFVDNDAGELNTRLTDYSKINDSGIKLGHEFPQSITTFSGAFTI	205	
Qy	244	AVYVKWIALICSSITVALVLTMGGSOPFIKYSKSLDSYGAGCTVAEEVISIRNATA	303	
Db	206	GPISQMKLTLLVLAASPLGLSSAMMAKVLFSTPKKEIQAVAKAGAFAEVLAIITFYIA	265	
Qy	304	FGTQOKLKROYEVHLDEAEKNWGTKNOIYWGMFGIMFGIMNGYGIQFMWSRELVDGAV	363	
Db	266	FGGQKKELETRYKNLEBAKRGVICKAITPANISIGLAYLVASYALAFMYGTSVLASNEY	325	
Qy	364	DVGDILLTVLMAILIGSFSLGNVSPAQAFTNAVAAAKIFGTIDROSPLDPSPNEGKTLD	423	
Db	326	SIGVLTVFEFSILTGTEIGHLAPNIIEFANRGAAYEIFKIIDNEPDISDSTGHKRPD	385	
Qy	424	HPEGHIELRNKHITPESREPYVMEDVLSMPAGKITLAVGSSGSKSVGVVERFYMP	483	
Db	386	SIMGMLERKNYENPSPSEVAILKGLNLKTVSGGTVALVANGSGGSKSTVOQLDRLPX	445	
Qy	484	VRGVITLLDGHDIKDLMLRLROOISLVSOEPFLFTTYIKNRHGLIGTRYENSESEDKVR	543	
Db	446	IEGEVSIIDGQDIRTINVRYLRREITGVSOEPLVFATTIAENIRYC-----RENVMIDE--	498	
Qy	544	ELIENAAKMANAHDFITALPEGYEITNVGORGFLTSGGQKORIAIARAAYSDBPKILLDEA	603	
Db	499	--IEKAVEANAAYDIIMKLPHKFDPLVGERGQOLSGQKORIAIARALVRNPKILLDEA	556	
Qy	604	TSAUDTKEBGVQAOLERAABGRRTIVIAHRSTIKTLANTIVLVNKGRIABOGTHDELYD	663	
Db	557	TSALDTSEBAYQAOLDKARKBGRRTIVIAHRISTVRNADVIAIGPDGGVIVEGNGNHELK	616	
Qy	664	RGGAYRKILEAQ-RINE-----QKEDALEDAEDAEULTNADIKIKITASSASSDDLQKPTT	718	
Db	617	EKGIFYKILVMOQTGRNEIEIPGNNAVESQSDIGASELTSESASPILRRSIRRSHRQQ	676	
Qy	719	IDRGTGHSVSSALLSKRPETTPRYSLMTLIKETAFASENREPEIPYMILCLVFSVLAGGQ	778	
Db	677	ERRRSSKDEDVD-----EDVPVVSFWQILK----LNISMPPYLAVGVLCALVINGCIO	723	

QY	779	PTQVAVAKKAIISTLSPESQYSKLRDADPMASLMEFVVGIIIOETQSTNGAAFAVACSRL	838
Db	724	PVALVESKIVGVPFSRDDHETQOR-NCNLFSLFELVMMISFVIFYFQGFPGACGTL	782
QY	839	IRARASTAFETILRODIAEPDKENSTGALTSFLSTETKHLSGSVGLTGLMTSTLG	898
Db	783	TKRLRYWVFKSMRLDPTSMFDDHKNTGSLTRRLASDASNVKANGSRLAVYTOQVANLG	842
QY	899	AALIIAIAI--GKMLAYCISVYPVLLACGFRTFYRLAOFQSRSLATEGSANFACETS	956
Db	843	TGIIISLVLYVGMQTLTLVVIIPLVIGGIIEMKLLSGQALDKKKEDEISGKATEAIE	902
QY	957	SIRTVASITREBRWVEIYHAOLDAGRTSLISVLRSSLLYASSQALYFECVALGFWYGT	1016
Db	903	NFRVYSLTREQAFENMYAQSLOIPRNLKKAHVIGTIFATQAMIFYTAACFBGAY	962
QY	1017	LGHEHYDIFRFVCESEILFGASAGTVEFSPADMGKAKNAAEFRLLFDRKPDQDMS	1076
Db	963	LVARELTFENWMLVFSAVVFGAMAGNTSSFPADYAKAKVASHIIIGIERPEIDXS	1022
QY	1077	EEGRKLETVGELIEFRVHRVRYTPREDOVYLGDLTVYRPGYVALVNPSCGKSTIAL	1136
Db	1023	TEGAKPWLGEJNVKFNKVPKRYTPRINIPVLGLSEYKAGGTLLVYSSGGKSTVQL	1082
QY	1137	LEREYDAIASIIVYDGKDISKLINSRFSFLSVSEPTLYOGTITKENILGIVEDPVE	1196
Db	1083	LERVYNNMAGTVFLDKEIKQLVQCVRA-LGIVSEPLTFPCSTIAENAVGDSRVYSH	1141
QY	1197	EELIKACKDANIYDFIWSLDEGPNIVYSGKGMISGGOKQVAVARALLRDPKILLDEA	1256
Db	1142	EELIYARARENHOETIDSPLEKNTYVGGKGTOLSSGGOKQRIARAVARVOPHILLDEA	1201
QY	1257	TSALDSESEKVOAALDAARGTTTAAVHRISTOKAVIYVFPQGVESGHSSELYQ	1316
Db	1202	TSALDSESEKVOBALDKAREGTCVIAHRUSTIOMADLIVIONGVKEGHQOULLA	1261
QY	1317	KKGRVYELV 1325	
Db	1262	OKGIYFSMV 1270	
RESULT 8			
MDR2_MOUSE			
ID	MDR2_MOUSE	STANDARD;	PRT; 1276 AA.
AC	P21440;		
DT	01-MAY-1991 (Rel. 18, Created)		
DT	01-MAY-1991 (Rel. 18, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Multidrug resistance protein 2 (P-glycoprotein 2).		
GN	ABCB4 OR PGY2 OR PGY-2 OR MDR2.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RA	SEQUENCE FROM N.A.		
RP	MEDLINE=88302195; Pubmed=3405218;		
RP	Gros P., Raymond M., Bell J., Housman D.;		
RT	"Cloning and characterization of a second member of the mouse mdr		
RT	gene family.";		
RT	Mol. Cell. Biol. 8:2770-2778(1988).		
RN	[2]		
RP	SEQUENCE OF 1-23 FROM N.A.		
RC	STRAIN-BALB/C;		
RA	Kirschner L.S., Horwitz S.B.;		
RL	Submitted (DEC-1991) to the EMBL/Genbank/DBJ databases.		
CC	-1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED		
CC	DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. MOUSE MDR2 IS		
CC	NOT CAPABLE OF CONFERRING DRUG RESISTANCE.		
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-1- MISCELLANEOUS: IN MOUSE THE MDR GENE FAMILY INCLUDES THREE OR MORE		
CC	RELATED BUT DISTINCT CELLULAR GENES.		

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J03398; AAA39516.1; -;
 DR EMBL: M74151; AAA39515.1; -;
 DR PIR: A30409; DVMS2.
 DR HSSP: P13569; INBD.
 DR MGI: P13569; Abcb4.
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABC_transport.
 DR Pfam: PF00005; ABC_tran; 2.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
 KW Multigene family.
 KM
 FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 52 75 POTENTIAL.
 FT TRANSMEM 119 139 POTENTIAL.
 FT TRANSMEM 188 208 POTENTIAL.
 FT TRANSMEM 215 235 POTENTIAL.
 FT TRANSMEM 296 316 POTENTIAL.
 FT TRANSMEM 325 345 POTENTIAL.
 FT DOMAIN 346 707 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 708 728 POTENTIAL.
 FT TRANSMEM 753 773 POTENTIAL.
 FT TRANSMEM 829 849 POTENTIAL.
 FT TRANSMEM 850 870 POTENTIAL.
 FT TRANSMEM 933 953 POTENTIAL.
 FT TRANSMEM 970 990 POTENTIAL.
 FT DOMAIN 991 1276 CYTOPLASMIC (POTENTIAL).
 FT NP_BIND 426 433 ATP (BY SIMILARITY).
 FT NP_BIND 1066 1073 ATP (BY SIMILARITY).
 FT REPEAT 636 1276
 FT SEQUENCE 1276 AA; 140332 MW; A6C38DCD2C118EEF CRC64;
 Query Match 35.58; Score 2387; DB 1; Length 1276;
 Best Local Similarity 42.18; Pred. No. 4.8e-133;
 Matches 534; Conservative 212; Mismatches 458; Indels 64; Gaps 13;
 QY 97 KVN-ISPFGIMRYATKMDIILIVISTICATAAASSTFORIML-YQISYDEYD----- 146
 DB 33 KVLILGLITLFRISDMQDKFLMFLGLTLMATLHSGSLPLMIVGEMTDKFTVDNTGNSLP 92
 QY 147 -----ELTRNLVEFYLLIGIEFYVYSTVGFIYTGHEATOKIREYYLE 190
 DB 93 VNFSLMLNGLRIIEEMETRYAYVYSGLGCVLVAAYIQVSEFWTLAAGROIKIKRQKFPH 152
 QY 191 SIIKRNIGYDKLGAGVYTRITADTDLIODGISEKGLTLLATATVYTAIYVYWK 250
 DB 153 AIIROEGWMDIRGTETLNRITDVSRISEIGIDKGMFEFOALATFPAFIVGFIWGMK 212
 QY 251 IALICSTYVALVLTMGGSOFIIRYKSKISDYAGAGTAAEEVYSSIRNATAGTQDKL 310
 DB 213 LTVIYMAISPIIGISTVYMAKILSTFSDKELAAVAKGAAVEEPGAIKRVIAFGGONKE 272
 QY 311 AKQEVHLDEAEKWKGTQNIIVGMFGAMFLMYSNGTGFWMGSRPLVDGADVGDILT 370
 DB 273 LERQKHLERNAKKIGIKKAIISANISMGIAFLIYASVALAFVGVSTLVISREYIGNAMT 332
 QY 371 VLAIIIGSSISGVSNAAAFNAVAAKIRGTIDROSLDIPYISNEGKTLHFBEGHIE 430
 DB 333 VFSIILIGASVGOAACPICIDAFNARGAAVYIEDIIDNNPKIDFSERGHKPDINKGLE 392

QY 431 LRVKHIYPSREYTVMAEDVSLSPAGKTALVYSSGSGKSTVGLYFERFMPYRGTVLL 490
 DB 393 FSDVHFETPSRANKIKILKGLNKLAKKSGQYALVNSGSGSTYQILQRLDLPREGIISI 452
 QY 491 DGHIDKDLNRWLNRQOISLVSQEPVLFGTYYKIRIRGLIGTKYENSEDKVRELIENAA 550
 DB 453 DGDIRNFNRVCLREIIGVVSQEPVLFSTIAENIRFG-----RGVNTMDE---IKAV 503
 QY 551 KMAAHDFITALPBGVETNNGORFLLSGOKQALAIARAVSPKILLDEATSALDTK 610
 DB 504 KEANAYDFIMKLPGKFPTLVGDGRGAQLSGQORAIARLVNPKILLDEATSALDTE 563
 QY 611 SEGVAQALBRAAEGRTTIVIAHRLSTIKTAHNIVLVNKGIAOGTHDELVDNGAYRK 670
 DB 564 SEAVQAALDKARGRRTTIVIAHRLSTIRNADVLAGEDDGYIVQGSHELMKREGIYFR 623
 QY 671 LVEAQRIINEQKADALDADAEDELTNADIAKIKTAYSSASSDLDGKPTTIDRTGHSVSS 730
 DB 624 LVNNGTASQILSEFEVELSDEKAGDAP-----NGMKARIFRNSTKSKLS 672
 QY 731 AILSKRPER-----TPKSLMTLLKFVASFNRPETPLMILIGVEYVLAGGQPTQAVL 784
 DB 673 PHQRLDEETNEIDANVPVPVSLKVLK---LNKTEPYPVGVCAIANGALQPARSII 728
 QY 785 YAKAISTLSIPESQYSKLRHDADFMVSMFVYVGIOTOSTNGAFAVCSERLIRARS 844
 DB 729 LSEMIATLFG-PGDDAVK-QQCKNMFSLVFLGLGLVSLFTPLGCTFGTKAGCEIITLRS 786
 QY 845 TAFRTILRODIAEFDEKENSTGALTSLSTETKHLGVSQVLTGTLMTSTTIGAAITIA 904
 DB 787 MAFKAMLRQMSWPDHDKNSTGALSTRLATDAQVATGATGKLLAIQNTANLGTGIIIS 846
 QY 905 LAIGMKLALVCISVVPVILACGFRTFMIAQFORSKLATEGSNFCEATSTRYASL 964
 DB 847 FIYGMQTLTLLSVPPVIAVAGIVEMKMLAGNAKRDKEKMAAKIATEIENRTVYSL 906
 QY 965 TREDDVEIYHAQDADAGRTSLISLRSSLLYASSQALVFECVALGFWYGGTLL--GHHE 1022
 DB 907 TOEKKEFSMYEKLHGYRNSVRKAHLYGTIFSLSQAFMFTSYAGCTRFSGSYLVNCHMR 966
 QY 1023 Y-DIREFVCFSEILFQAQSGATVFSFAPDMGKAKNAAEFRRLFDKRPQIDNMSEGEK 1081
 DB 967 FKDY---ILVFSALVLAVALGHAASSFAPDPAKAKLSAAYLFLSEFQPLIDSYSGGLW 1023
 QY 1082 LFTYEGEIERRNHFRPTPEQPVLAGDLTVAPGQYVALVSGGCKSTTIALLERFY 1141
 DB 1024 PDKEGSGVTENEVFNPTIRANVPVLOGLSLEVKKQTLALVSGGCKSTVQDLLEFY 1083
 QY 1142 DALAGSLIVGKDSIKINISYRSFLSVSOEPTLVYGTIKENTLLGIVEDDVEEFLIK 1201
 DB 1084 DPMAGSVLLDGOEAKKILANQMLRQLOGISOEPFLFPCSTIAENIAYDNSRVPHDITVR 1143
 QY 1202 ACKDANTYDEIMSLPEGFNTVYSGSGMLSGGOKORVAIARALRDPIKILLDEATSALD 1261
 DB 1144 AAKRANIHPIETLTPQYNTVRVGDGKGTOLSGGQORAIARALIRQPRVILLDEATSALD 1203
 QY 1262 SESKVVQAALDAARGRITIAVAHRLSTOKADVITYVPOGKIVESGTSSELVQKGRY 1321
 DB 1204 TESKVVQEAALDKRREGRTCIIVAHRLSTIONMDLIYIENGKVKKEGTHOOLLAQGIY 1263
 QY 1322 YELVNIQS 1329
 DB 1264 FSWNIQA 1271
 RESULT 9
 MDR3_HUMAN
 ID MDR3_HUMAN STANDARD; PRT; 1279 AA.
 AC P21439;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN ABCB4 OR PGY3 OR MDR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89138016; PubMed=2906314;
RA van der Bliek A.M., Koelman P.M., Schneider C., Borst P.;
RT "Sequence of mdr3 cDNA encoding a human P-glycoprotein.";
RL Gene 71:401-411(1988).
RN [2]
RP SEQUENCE OF 856-1279 FROM N.A.
RX MEDLINE=88111519; PubMed=2892668;
RA van der Bliek A.M., Baas F., ten Houte de Lange T., Koelman P.M.,
RT "The human mdr3 gene encodes a novel P-glycoprotein homologue and
RT gives rise to alternatively spliced mRNAs in liver.";
RL EMBO J. 6:3325-3331(1987).
RN [3]
RP GENE STRUCTURE.
RX MEDLINE=91161629; PubMed=2002063;
RA Lincke C.R., Smit J.J.M., van der Velde-Koerts T., Borst P.;
RT "Structure of the human MDR3 gene and physical mapping of the human
RT MDR locus.";
RL J. Biol. Chem. 266:5303-5310(1991).
CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS. HUMAN MDR3 IS NOT
CC CAPABLE OF CONFERRING DRUG RESISTANCE. MEDIATES THE TRANSLOCATION
CC OF PHOSPHATIDYLCHOLINE ACROSS THE CANALICULAR MEMBRANE OF THE
CC HEPATOCYTE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DISEASE: DEFECTS IN ABCB4 ARE A CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS TYPE IIT (PFIC), A FORM OF AUTOSOMAL
CC RECESSIVE LIVER DISORDERS, CHARACTERIZED BY EARLY ONSET OF
CC CHOLESTASIS THAT PROGRESSES TO CIRRHOSIS AND LIVER FAILURE BEFORE
CC ADULTHOOD.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23234; AAA36207.1; -;
DR EMBL: X06181; CAA29547.1; -;
DR PIR: J50051; DVHJ3.
DR PIR: A42213; A42213.
DR HSSP: P13569; INBD.
DR Genem: HGNC:45; ABCB4.
DR MIM: 171060; -;
DR MIM: 602347; -;
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001440; ABCtranspfltm.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD00006; ABC_transporter; 2.
DR SMART: SMO0382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 123 78 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
FT TRANSMEM 192 211 POTENTIAL.
FT TRANSMEM 216 235 POTENTIAL.
FT TRANSMEM 301 320 POTENTIAL.
FT TRANSMEM 336 354 POTENTIAL.

FT DOMAIN 355 711 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 712 732 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 832 851 POTENTIAL.
FT TRANSMEM 854 873 POTENTIAL.
FT TRANSMEM 937 956 POTENTIAL.
FT TRANSMEM 976 993 POTENTIAL.
FT DOMAIN 994 1279 POTENTIAL.
FT NE_BIND 429 436 POTENTIAL.
FT NE_BIND 1069 1076 POTENTIAL.
FT REPEAT 1 640 POTENTIAL.
FT REPEAT 641 1279 POTENTIAL.
FT CARBOHYD 91 91 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1093 1093 V -> VEVDFGRQ (IN REF. 2).
SQ SEQUENCE 1279 AA; 140682 MW; 3D58C9858CB06087 CRC64;
Query Match 35 48; Score 2379.5; DB 1; Length 1279;
Best Local Similarity 40.88; Pred. No. 1,3e-132;
Matches 540; Conservativity 231; Mismatches 466; Indels 87; Gaps 17;
QY 42 DLAPSTTATPADKKEHPKSSNNNAVSEVDALIAHLPEDROVLTQLEETVNIS 101
DB 2 DLEAAKNGRA-----WRPTSAEGDFELGIS-----KORRKYT---VAM-IG 41
QY 102 FGLMRYATKMDILIAVISTICAIAAASFQRTLM-YQISYDFYD----- 146
DB 42 VTLFTRSPDQDLFLMSIGLITAIAGSGGLPLMIYFGEKTDKFDVTAQNSFPVNFSS 101
QY 147 -----ELKKNLYFVYLGIGEFYVYVSTGVFIYTGAEATQKIREYILESTLRON 196
DB 102 LNPGLILEEMRYVAYVYSGAGVLAAYIOVSPWTLAAGROIKRQKFEHALROR 161
QY 197 IGYFDKLAGEVYTRITADTNLIODISEKVGTLTALATFVAFIAYVKKWKLKLS 256
DB 162 IGFVDINDTTELMTFRITLDISKISEGIDKVGKFPQAVATFPFGFIVGRKRLVLIM 221
QY 257 STVALVLTNGSGOPFIITYSKSLDSYAGAGTVAEYISSIRNATFSTQDLAKQYEV 316
DB 222 AISPIGLSAAWAKILSAFSDKEILAAVAKAGAVAEALGAIPTVAFGGONKELERYOK 281
QY 317 HIDEAKKWKGNQIVYFPMIGAMFGLMYSNYGLGFNGSRFLVDGAVDQDILTVLML 376
DB 282 HLENAKEITIKKAIKSNISMGIAFLDLIVASYALAFVGGTSLVYSKYTTGGMATVFSSIL 341
QY 377 ISFSISLGNVSPNAQFTNVAANAIFGTIDROSPLDPSNBEKTLDFHEGHTELNVKH 436
DB 342 IGAFSVGAAPCIDAFANRGAAYVFFDILNNPKIDSPSEGRCHKPDSIKGNLEFNDVAF 401
QY 437 IYSPREYIVMEDVSLSPAGKTTALVSPSGSKSTVGLVERFYMPVPGTVLLDGHDIK 496
DB 402 STPSRANVILKGLNKLKYGQTVLVSSGCGKSTTVQLRIQLDPDDEGTINIDGDIR 461
QY 497 DLNLRLROOISLVSOEPLVFGTIIYKTRHGLIGIKKYNESEDKVRELLENAAKMANH 556
DB 462 NKNVNTLRITIGVSOEPLVSTTIAENICYG-----KGNVMTDEKK---AVEANAY 512
QY 557 DFTALPEGYETNVGSGRGLVSGGOKORIAIARAVYSDPKILLDEATSAIDTKSEGVYQ 616
DB 513 EFLMKLPQKFDILVERGQSGGOKORIAIARALVYRNPKILLDEATSAIDTKSEAEVQ 572
QY 617 AALEPAAEGRTTIVIAHRLSTIKTANIVLVNGRIAGQTHDELVDNRGAVRKLVEAQR 676
DB 573 AALDRAREGRTTIVIAHRLSTVRNDVIAGFEDVIVEGSHSELKKKEGVFKLVNQT 632
QY 677 INEQKEADALEADADLTNADIAIKITASSASSDLDGKPTTIDRTGHRKYSVAIILSKR 736
DB 633 GSGQIOSEFE-----LNDEKATRMAP-----NWKSKSLFRHSSTOKNLKNSQMCOK 679
QY 737 PPE-----TPKYSILMTLKFVASFNRPEIPMLIGIVSVYLAGGQOPQAVLYAKA 788
DB 680 SLIDVETDGLKANVPVSLFKYLK-----LNKTEWPFVVGVCALANGGLQPAFVSIFSEI 735


```

DB 554 DEFTSALDTSSEAEVQALADKAREGRTIYIAHRLSTVRADYIAGDEGYIEQSHSE 613
QY 661 LVDRGAYRKLVEAQRINECKEADADADADLTNADIAKIKITASSASSDLDGKPTTID 720
DB 614 LIKEGTYFELVLMQTSQSIILSEEF-----VELSDERAKAGVAP--NGMKARIF 662
QY 721 RTGTHKSVSA-----ILSKRPETTKYSIMTLTKKLVASPNREIYMLIGLVSV 772
DB 663 RNTKSKSKSRHONRLDETVELDANVPVPSFLKYL---LNKTEMPEYFVVGITCAI 718
QY 773 LAGGQPTQAVLYAKAISTLSPESQSKLRHDAFWSLMEFVVGIIQFTQSTNGAFA 832
DB 719 ANGALPAPFIIISEMIAIIG-PGDDTVK-QQCNMFSLVFLGIVHSFTFLQGTG 776
QY 833 VCEBLIRARSTAFRTIIRODIAFPDKENSIGALTSFSTETKHLISGVSVTLGITIM 892
DB 777 KAGEIILTRLSMAFMALQDMSWEDKHNSGTALSTRLATDAVQAGATGRLALIAQ 836
QY 893 TSTLGAIIIALIGMKIALVCISVYPVLLAGGYRFVYLAQFOSKSLAEGSANFAC 952
DB 837 NTANLGTGIIISFTYGQTLTLISVPTAVAGIVEMKMLAGNAKDKKEMEAQRIAT 896
QY 953 EATSIRTVASLTRERDWEIYHAOLDAQRTSLISYLRSSLYASSQALVFCVALGFY 1012
DB 897 EAIENIRTVASLTRERDWEIYHAOLDAQRTSLISYLRSSLYASSQALVFCVALGFY 956
QY 1013 YGTLT--GHHEX-DIFRFVCSSEILFGAQSAGTVSPAPDNGKAKNAAEFRRLPDRK 1069
DB 957 FGSYLLVNGHMRKDY---ILVFSAYILGAVAGHASSFPDYAKAKLSAAYLFSEERQ 1013
QY 1070 POLDNMSSEGEKLEVEGEIEFNHVRYPTRPEOVLRLGLDITVRCGVVALVSGSGC 1129
DB 1014 PLIDTSRREGMDKREKSGSTFNEVFNPTFRANVPLQSLSEVKKQGLTAVGSSGCG 1073
QY 1130 KSTTIALLEFYDLAGSILVYDKDISKLINSRSLVSQEPLOYGTIKENILIGI 1189
DB 1074 KSTVOLLEREFYDPMAGTVLLDQGAOKLVQMLRGLVSGEPILFQDSIAKNATYGD 1133
QY 1190 VEDVPEEFLIRACRANAYIDFIMSLPEGFNYVSGKGLSGGQKORAVATARALIDPK 1249
DB 1134 NSRWVSQDIDIVRAKCANHPFETLPKRYETRVGDKGTOLSGQKORAIARALRQPR 1193
QY 1250 ILLDPAATSDSESEKVOAALDAARGRTIYAVAHRLSTOKAVIYVFDGKIVEGS 1309
DB 1194 VILLDPATSDLDESKEKVOEALDKAREGRTIYIAHRLSTIONADLIIVINDGKVENG 1253
QY 1310 THSELVQKRGRYEELVNLQS 1329
DB 1254 THQQLLAGKGYFSMVNIQA 1273

RESULT 11
MDR3_CRIGR STANDARD: PRT: 1281 AA.
AC P23174:
DF 01-NOV-1991 (Rel. 20, Created)
DF 01-NOV-1991 (Rel. 20, Last sequence update)
DF 16-OCT-2001 (Rel. 40, Last annotation update)
DE Multidrug resistance protein 3 (P-glycoprotein 3).
GN PGY3 OR PGY3.
OS Cricetus griseus (Chinese hamster).
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92135896; PubMed=1685679;
RA Endicott J.A., Sarangi F., Ling V.;
RT "Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
RL DNA Seq. 2:89-101(1991).

```

```

CC -1- FUNCTION: ENERGY-DEPENDENT EFFLUX PUMP RESPONSIBLE FOR DECREASED
CC DRUG ACCUMULATION IN MULTIDRUG-RESISTANT CELLS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- MISCELLANEOUS: PGP ISOFORMS DIFFER IN THEIR DRUG TRANSPORT
CC CAPABILITIES: PGP1 AND PGP2 CAN MEDIATE MDR, WHILE PGP3 APPARENTLY
CC CANNOT.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M60042; AAA68885.1; -.
DR HSSP: P13569; INBD.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABCtransprtTM.
DR Pfam: PF00005; ABC_tran; 2.
DR Pfam: PF00664; ABC_membrane; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat;
KW Multigene family.
FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 78 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 191 211 POTENTIAL.
FT TRANSMEM 218 238 POTENTIAL.
FT TRANSMEM 299 319 POTENTIAL.
FT TRANSMEM 328 348 POTENTIAL.
FT DOMAIN 349 712 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 713 733 POTENTIAL.
FT TRANSMEM 758 778 POTENTIAL.
FT TRANSMEM 834 854 POTENTIAL.
FT TRANSMEM 855 875 POTENTIAL.
FT TRANSMEM 938 958 POTENTIAL.
FT TRANSMEM 975 995 POTENTIAL.
FT DOMAIN 996 1281 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 429 436 ATP (POTENTIAL).
FT NP_BIND 1071 1078 ATP (POTENTIAL).
SQ SEQUENCE 1281 AA; 140866 MW; 2203BF61EBB29602 CRC64;

Query Match 35.1%; Score 2357; DB 1; Length 1281;
Best Local Similarity 41.3%; Pred. No. 2.9e-131;
Matches 526; Conservative 224; Mismatches 449; Indels 76; Gaps 13;

QY 97 KVN-ISFGLMRYATKMDILIMVITICAAIAAASFQRIIML-YQISYDEFY----- 145
DB 36 KYNLIGPLPLFRSDQMDLFLMIGTIMAIAHSGSLPLMIYFGEKTDVFNNAAGFSLP 95
QY 146 -----DELTKNLYEYVYLGIGEFVTVVSTVGFIYTGHAHQKIREYLE 190
DB 96 VNFSLIMINPGRILREEMKRYAVYISGLGGVLYAVYIQVSWTLAAGQIKKIRQNFH 155
QY 191 SITRONIGFEDKAGEVYTRITADINLQDGSSEVGLTALATFTVAFTIAYKYTK 250
DB 156 ALTRQDMGFMDIKGTETLWRLDIDSKISEGIGDVGFQAVAFVFGFTVGFRGK 215
QY 251 LALICSTIVALVLTWGGSGQFLIKSKSLDSYGAGTVAEVISSINATAFGQDRL 310
DB 216 LTLVMAISPIILGSAVAAKILSTFSDKEIAAYAKAGAVAEALGATVYAFGQNK 275
QY 311 AKOYEYHLDIAEKWGKTKNOIVGFMIGANFGLMTSYNGLGFMMGSRFLVDGAVDGLT 370
DB 276 LERYOKHLENARKIGIKKAISANISWGIALLIYASYALAFYVGSPLVSKETYGNAWT 335
QY 371 VMALIGSFSIGNVSPNNAQFTNVAANAATFETIDRQSPDIPYSNEKTIIDHEGHLE 430

```

```

Db 336 VEPISILIGAFSGOAPCIDAFANARGAAYIEDIIDNNKIDFSERGHKPKDSIKGNLD 395
QY 431 LRRNKHLYPSRPEYTMEDVSLSPACKTATLVSPSGSGKSTVGLVERFYMPRGTVLL 490
Db 396 FSDVHPSYPRANIKIKLKGTLKQVSOQYALVGNSSCGTTLTQILQRLYDPEGIISI 455
QY 491 DGHDIKDLNRLMRLQOISLVSOEVLFGTTIYKNIRHGLIGTKYENESDKVRELIENAA 550
Db 456 DGODIRNFNRVRLREITIGVSOEVLFTSTIAENIRFG-----RGVNTMEIKKAV 506
QY 551 KMANADFTALBEGYTNGVSGFLLSGGOKRIALARAVIDPKLLIDDEATSAIDTK 610
Db 507 KEANAAEFIMKLPOKEFTLVGERGAQISGGOKRIALARLVNPKILLIDDEATSAIDTE 566
QY 611 SEGVQALERRAAGRTTYIARHLSITKTAHNVLVYNCKIAKBOGHDELVDGAYRK 670
Db 567 SEAEVQALDKAREGRTTYIARHLSITKTAHNVLVYNCKIAKBOGHDELVDGAYRK 626
QY 671 LVE-----AQRINEOKBADLEDADADLTNADIATKRTASASSDDLGRPTTIDRTGTH 725
Db 627 LVNMQTSGSOLSOEFEEVLESEKADGMTP-----NGMKSHIFRNSTK 670
QY 726 KVSUSA-----ILSKRPETTPKSLMTLTKFVASFNRPEIPYMLIGVSVLAGGG 777
Db 671 KSLKSSAHHRRLDVEDDELDAVNPVSLKVLK---LKTETPVYFVGVCALVNGAL 726
QY 778 OPTQAVVYAKAISTLSPEQSOKLRHDADFWSLMFVVGIIORTITOSTNGAFAVCSER 837
Db 727 QPASTILSEMAIFG-PGDDANK-QOKMLFSLVFGIGLVSTFTFLOGFTTGKAGEI 784
QY 838 LIRRASTAFRTLLQDIAFEDKEENSTGALJTSFLSTETKHLSCVSGVTGTLTMTSTLL 897
Db 785 LTRLRSMARFAMLRQMSWFDYKNSTGALSTRLATRAVOGATGTRILLIIONTANL 844
QY 898 GAAITIALAGMKIALYCSIVPVLLACGFYRFLMLAOPSRKSLAEGSNFCESTSS 957
Db 845 GTGIIISIFYGMOTLLTLLSVPIAVSGIVEMKMLAGNAKRDKALEAGKTAIEN 904
QY 958 IRTVASLSTRERDWEIYHAOLDAGRTSLISVLSLSSLYASSQALVFECVLAGFWYGTLL 1017
Db 905 IRRVVSILTORKEPESMVEKLEHPRYNSGVMAHITGTFISQAFMFTSAGCFRGAYL 964
QY 1018 L--GHHEY-DIFRFVCFSEILFGAQSAGVTFSPADPMGRKAKNAAEFRLLDRKPOIDN 1074
Db 965 IVNGHMFRRDY---ILVFSALVFGAVALGHASSFAPDARAKLSAHLFSEFQPLIDS 1021
QY 1075 WSEGEKLEVEGEIERANHERFTRPREQPVLRGLDLYKPGQYVALVGPBSGGKSTTI 1134
Db 1022 YSGGLMPDFEGSVTFNEVFNPTTANMPVLQGLSLEVKKGOTLALVSGSGGKSTVV 1081
QY 1135 ALLERFYDAJAGSLVDGDKISKLINSYRSFLSVSOEPTLYOGTITKENILLGIVEDV 1194
Db 1082 QLLERFDPAGVYLLDGGQAKKINIQMLAQLGIVSOEPTLYOGTITKENILLGIVEDV 1141
QY 1195 PEEFLKACKDANITYDIMSLEPGENTVSGSGKMLSGGOKORAVIARALLDRPKILLD 1254
Db 1142 SOBEIVAAANAHNIFETLPQKRYKRGVKGFTQLSGGOKORAIARRALLDRPKILLD 1201
QY 1255 EATSALDSESEKVVQALDAAARGRTTIANHRLSTQKADVITYVPOGKIVESGISEL 1314
Db 1202 EATSALDSESEKVVQALDAAARGRTTIANHRLSTQKADVITYVPOGKIVESGISEL 1261
QY 1315 VOKGRYELVNLOS 1329
Db 1262 LAQKGIYFSMNTQA 1276

```

```

RESULT 12
AB11_HUMAN STANDARD; PRT; 1321 AA.
AC 095342: 09JUN82:
DF 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11).
GN ABCB11 OR BSEP.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND VARIANTS PFC2 GLY-297; GLU-461; GLY-482;
RP ARG-982; CYS-1153 AND GLN-1268.
RX MEDLINE=99021377; PubMed=9806540;
RA Strautnieks S.S., Bull L.N., Knisely A.S., Kocoshis S.A., Dahl N.,
RA Ansell H., Sokal E., Dahan K., Childs S., Ling V., Tanner M.S.,
RA Kagalwalla A.F., Nemeth A., Pawlowska J., Baker A., Miel-Vergani G.,
RA Feghali N.B., Gardner R.M., Thompson R.J.;
RT "A gene encoding a liver-specific ABC transporter is mutated in
RT progressive familial intrahepatic cholestasis."
RL Nat. Genet. 20:233-238(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Mol O., Hooiveld G.J.E.J., Jansen P.L.M., Muller M.;
RT "Cellular localization and functional characterization of the human
RT bile salt export pump (BSEP).";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP VARIANT ALA-444.
RX MEDLINE=21686803; PubMed=11829140;
RA Saito S., Iida A., Sekine A., Miura Y., Ogawa C., Kawauchi S.,
RA Higuchi S., Nakamura Y.;
RT "Three hundred twenty-six genetic variations in genes encoding nine
RT members of ATP-binding cassette, subfamily B (ABCB/MDR/TAP), in the
RT Japanese population."
RL J. Hum. Genet. 47:38-50(2002).
RN [4]
RP VARIANTS INTRAHEPATIC CHOLESTASIS LEU-284 AND ASP-1004.
RX MEDLINE=21674594; PubMed=11815775;
RA Chen H.-L., Chang P.-S., Hsu H.-C., Ni Y.-H., Hsu H.-Y., Lee J.-H.,
RA Jiang Y.-M., Shau W.-Y., Chang M.-H.;
RT "PFC1 and BSEP defects in Taiwanese patients with chronic intrahepatic
RT cholestasis with low gamma-glutamyltranspeptidase levels."
RL J. Pediatr. 140:119-124(2002).
CC -I- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
CC INTO THE CANALICULUS OF HEPATOCYTES.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
CC SITU (BY SIMILARITY).
CC -I- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC -I- DISEASE: DEFECTS IN ABCB11 ARE THE CAUSE OF PROGRESSIVE FAMILIAL
CC INTRAHEPATIC CHOLESTASIS 2 (PFC2), AN INHERITED LIVER DISEASE OF
CC CHILDHOOD. PFC2 IS CHARACTERIZED BY CHOLESTASIS AND NORMAL SERUM
CC GAMMA-GUTAMYLTRANSFERASE ACTIVITY. DEFECTS IN ABCB11 ARE ALSO
CC FOUND IN CASES OF CHRONIC INTRAHEPATIC CHOLESTASIS WITHOUT OBVIOUS
CC FAMILIAL HISTORY OF CHRONIC LIVER DISEASE.
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF091582; AAC7455.1; -.
CC EMBL; AF136523; AAD28285.1; -.
CC Genev; HGNC:42; ABCB11.
CC MIM; 603201; -.
CC MIM; 601847; -.
CC InterPro; IPR003593; AAA_ATPase.

```

DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001140; ABC_transport_TM.
DR Pfam: PF000605; ABC_tran. 2.
DR Pfam: PF00064; ABC_membrane. 2.
DR ProDom: PD000006; ABC_transport. 2.
DR SMART: SM00382; AAA. 2.
DR ProSITE: PS00211; ABC_TRANSPORTER. 1.
KW ATP-binding; Transmembrane; Transport; Disease mutation; Polymorphism.
FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 84 147 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 169 215 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 237 240 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 241 261 POTENTIAL.
FT TRANSMEM 262 319 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 320 340 POTENTIAL.
FT TRANSMEM 341 353 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 354 374 POTENTIAL.
FT TRANSMEM 375 755 POTENTIAL.
FT TRANSMEM 756 776 POTENTIAL.
FT TRANSMEM 777 794 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 795 815 POTENTIAL.
FT TRANSMEM 816 869 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 870 890 POTENTIAL.
FT TRANSMEM 891 911 POTENTIAL.
FT TRANSMEM 912 979 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 980 1000 POTENTIAL.
FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1012 1032 POTENTIAL.
FT TRANSMEM 1033 1321 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 455 462 ATP (POTENTIAL).
FT NP_BIND 1113 1120 ATP (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 284 284 V -> L (IN CHRONIC INTRAHEPATIC CHOLESTASIS).
FT VARIANT 297 297 /FTID-VAR_013332.
FT VARIANT 444 444 E -> G (IN PEIC2).
FT VARIANT 444 444 /FTID-VAR_010271.
FT VARIANT 444 444 V -> A.
FT VARIANT 461 461 /FTID-VAR_013333.
FT VARIANT 461 461 K -> E (IN PEIC2).
FT VARIANT 482 482 /FTID-VAR_013334.
FT VARIANT 482 482 D -> G (IN PEIC2).
FT VARIANT 982 982 /FTID-VAR_013335.
FT VARIANT 982 982 G -> R (IN PEIC2).
FT VARIANT 1004 1004 /FTID-VAR_013336.
FT VARIANT 1004 1004 G -> D (IN CHRONIC INTRAHEPATIC CHOLESTASIS).
FT VARIANT 1153 1153 /FTID-VAR_013337.
FT VARIANT 1153 1153 R -> C (IN PEIC2).
FT VARIANT 1268 1268 /FTID-VAR_013338.
FT VARIANT 1268 1268 R -> Q (IN PEIC2).
FT CONFLICT 339 339 /FTID-VAR_013339.
FT CONFLICT 339 339 V -> L (IN REF. 2).
SQ SEQUENCE 1321 AA; 146392 MW; D44ACACC48DEA371 CRC64;
Query Match 34.6%; Score 2328; DB 1; Length 1321;
Best Local Similarity 38.7%; Pred. No. 1.5e-129;
Matches 508; Conservative 237; Mismatches 469; Indels 100; Gaps 14;

QY 161 GEPYVYVSVGVFIYGEHATQKIREYYLIESILRONIGYEDKACAGVTRITADNLQ 220
DB 151 AVLTITGICQEVYIAAARQIQKRRKFRIMMELGWEDCSVSGELNTRFSDIDKIN 210
QY 221 DGISEKVGTLTALVFAVFIILAYVYKMLALICSSITVALYVTMGGSOFI-----IKY 276
DB 211 DALADQALALIQNTMTISGIFLIGFFGKMLTY-----IISVPLIGATIGLSVKF 266
QY 277 SKSLDSYGAGGYAAEVISSIRNATAFGQDILAKQYEHLEDAEKMGTRKNOIVGMFI 336
DB 267 TDVELKAYARAGVAVADEVISMRTVAFAFGKEKEVEYERKNVFAQWGRKGIYMGFFT 326
QY 337 GAFGLMYNXYGCGFWMGSRFLVD-GAVDVGDIITVIMALLISFSLGNVSPNAQAFNA 395
DB 327 GFWCLLEFLCYAAVFWGSLVYLDEGEYTGTVQVLEFLSYVALNIGNASPCLAEAFATG 386
QY 396 VAAAKIFGTIDRQSLDPYNSNEKTLDHDEGHLELRNHYIPSRREVVMEDVSLMP 455
DB 387 RAATSTIFETIDRKPIIDCSBEGYKIDRIKGELEPRNVTFFHPSRPEVALIDNLNVIK 446
QY 456 AGRTTALVPSGSGKSTVGLVERFYMPVPGTVLLDGHDKLNLRLQOISLVSOEPY 515
DB 447 PGEMTALVPSGSGKSTALQIQRFYDPCGAVTVGHDIRSLNIQMLRQIGIVQEPY 506
QY 516 LPGETTYKINRHLGIKTYENESDKVRELIEAAKAANAHDITVALPGSETNVOGRGF 575
DB 507 LFSTTIAENIRYG-----REDATMEDIVQAANKANAYNIMPLPOQFDPLVGEGG 557
QY 576 LLSGQKORAIARAVYSDPKILLDEATSAULTKSGVQAALERAABERTIYVAHRL 635
DB 558 QMGSGQKORAIARALIRNKKILLDMATSALENSBAMQEVLSKIQHHTIYVAHRL 617
QY 636 STIKTAHNVIVLNGKIAEGTHDELVDRCGAYRKLVLEAORINEKEADLEADADLT 695
DB 618 STVRADTIIFGFHGAVERGTHEELERKGYFTVLQSQNQ---ALNEDIKDAT 673
QY 696 NADI-----AKITAS-SASDIDKRP--TTIDRTGHS--VSSAILSK 735
DB 674 EDMILARTSRGYSYDLSRISIRQSRKQSLSYVHEPLAVVHKSITYEDRDKDIPVQ 733
QY 736 RPETTPKYSILMTLKFVASFNRPEIPYMLIGLVFSLVAGGGQPTQAVYAKAISTSLP 795
DB 734 EHEEPAP-----VARILKFSAPEMPYMLVSGVGAANCTVPLAFLFSQLGTFPSLP 786
QY 796 ESQYSLKLRHADWLSLMEFYVGIIOFTOSTNGAFAVCSERILIRARSTAFETIIRODI 855
DB 787 DKE--RQRQINGVCLLEFVAMGCVSLFTQFLQYARAKSGELLTKLRKGFRAMLGQDI 844
QY 856 AFDKRENSGATLSLSTETKHLSSGVSYTGLTIIMSTITGLAATITLALIGWKALVC 915
DB 845 AMFDDLRSNPGALTTIRLADASQVGAAGSQIOMIVNSFTNVTVMILAFSFSWKLSTVI 904
QY 916 ISVVPVLLAGCFRYFMAFOFORSKLAVEGSANFACETSSIRTVASLIREDDWEIYH 975
DB 905 LCFPPPLASGATQTMILGFPASRDQALEMGQITNEALSIRTVAGIGKERRPTEALE 964
QY 976 AQLDAQRTSLISLVLSLLYASSQALVEFCVALGFWGTLGHEHYDIFREYVCFSEI 1035
DB 965 TELEKFPKTAIQKANIYGCFAFACIMFANSARYGYILNSGHLFSYFRTISAV 1024
QY 1036 LFAQASAGTVFAPDPMGAKNAAEFRRLFDKRPQIDMMSSEGEKLEFVEGIERNVH 1095
DB 1025 VLSATVLAGRAFSYTPPAKAKISAAFFQILLDRQPISYVNTAGEKMDNFQKIDVDC 1084
QY 1096 FRPPTAPQVPLRGDLTVFKPGQYVAVLPSPGSGKSTTALIERFDALAGSLVNGXKI 1155
DB 1085 FLYPSRPDQVNLGLSVSLSPGQTLAFVSSGCGKSTSIQLERFYDPDQGYMIDGHS 1144
QY 1156 SKLINYSRFSLSVSOEPTLYOGTITKENITLIGIVEDVPEEFELIKACDANIYFIMSL 1215
DB 1145 KKVNVQFLSNIGIYQSEVPLFACSIMDKIKGDNKELPMERYVIAAQAQDLHPVMSL 1204
QY 1216 PEGFNTVSGSKGMLSGGQKORVATARALLRDPKILLDEATSALESEKVVQAALDA 1275

DB 1205 PEKYTEWGSQSGSOLSRGEQRIALAAVYRDKILLDLDTATSLDSESKYQVALADKA 1264

OY 1276 ARGRTIAVAHRLSTIOKADVIYVPOGKIVESGTHSELVOKGRYEVLNLOS 1329

DB 1265 RESRTCTVAHRLSTIONADIIVNMAQGVIEGTHHELMAGAKYKLVTTGS 1318

RESULT 13

AB11_MOUSE STANDARD: PRT; 1321 AA.

AC 090Y30: 090ZEB: 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)

GN ABCB11 OR BSEP OR SPCP.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_Taxid=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=20076398; PubMed=10607905;

RT Green R.M., Hoda F., Ward K.L.;

RT "Molecular cloning and characterization of the murine bile salt export pump";

RL Gene 241:117-123(2000).

RN (2)

RP SEQUENCE OF 463-635 FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Liver;

RA Salkar R., Suchy F.J., Ananthanarayanan M.;

RT "Molecular cloning of mouse liver bile salt export pump (bsep).";

RT Submitted (SEP-1999) to the EMBL/GenBank/DBD databases.

CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS INTO THE CANALICULUS OF HEPATOCYTES.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN SITU.

CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES, EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN ATP BINDING CASSETTE (ABC) DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb.slb.ch).

CC -----

DR EMBL: AF133903; AAF14372.1; -

DR EMBL: AF186585; AAD56419.1; -

DR MGD: MGI:1351619; Abcb11.

DR InterPro: IPR003593; AAA_ATPase.

DR InterPro: IPR003439; ABC_transporter.

DR InterPro: IPR001140; ABCtransporterTM.

DR Pfam: PF00005; ABC_tran; 2.

DR Pfam: PF00664; ABC_membrane; 2.

DR ProDom: PD000006; ABC_membrane; 2.

DR SMART: SM00382; AAA; 2.

DR PROSITE: PS00211; ABC_TRANSPORTER; 1.

KW ATP-binding; Transmembrane; Transport.

FT DOMAIN 1 62 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 63 83 POTENTIAL.

FT DOMAIN 84 147 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 148 168 POTENTIAL.

FT DOMAIN 169 215 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 216 236 POTENTIAL.

FT DOMAIN 237 240 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 241 261 POTENTIAL.

FT DOMAIN 262 319 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 320 340 POTENTIAL.

FT DOMAIN 341 353 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 354 374 POTENTIAL.

FT TRANSMEM 375 755 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 756 776 POTENTIAL.

FT TRANSMEM 777 794 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 795 815 POTENTIAL.

FT TRANSMEM 816 869 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 870 890 POTENTIAL.

FT TRANSMEM 891 911 POTENTIAL.

FT TRANSMEM 912 979 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 980 1000 POTENTIAL.

FT TRANSMEM 1001 1011 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 1012 1032 POTENTIAL.

FT TRANSMEM 1033 1321 CYTOPLASMIC (POTENTIAL).

FT NP_BIND 455 462 ATP (POTENTIAL).

FT NP_BIND 1113 1120 ATP (POTENTIAL).

FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 481 481 L -> P (IN REF. 2).

FT CONFLICT 633 633 T -> V (IN REF. 2).

SQ SEQUENCE 1321 AA; 146675 MW; 15B5EBF15D32967 CMC64;

Query Match 33.4%; Score 2247; DB 1; Length 1321;

Best Local Similarity 38.1%; Pred. No. 9,4e-125;

Matches 502; Conservativity 246; Mismatches 456; Indels 112; Gaps 19;

OY 90 KTOLEEK-----VNISFGLMRVATKMDIIMVISTICATA--AASTQRMLOISDER 144

DB 31 KSRLODKKKKGARGVFELFRSSKDNLMFMGVCALLHGNAQGMIVFGILTDIF 90

OY 145 --YD-----ETKNVLF--VYLGIGE 162

DB 91 VERDIERGELISPGKVCNMNTIWINSPNOMNNGSCGLVDINSEVIFSGIYAGV 150

OY 163 FTVV--YVSTVGFIYTGHAHTQKIREYLLSILRONTGYFDKLGAEVTRITADTNLIQ 220

DB 151 AVLLIGFQIIRLWITTGABOIRKKRKYFRIRIMEIGFDCSVGLNRSFSDINKID 210

OY 221 DGISEKGLTLTATVTVAFTIAYVYKIALICSTIYALVLTMGGSQFT----IKY 276

DB 211 EAIADQALFLQRLSTALSGLLGFYGRKLTIV---ILAVSPLIGIGAIVGLSVAKF 266

OY 277 SKKSLDSYGAGVAAEVVSSIRNATAGTODKLAKQYVHLDEAEKRGKNOIVMGFMI 336

DB 267 TELELRKAYAKAGSTADEVLSSIRVAAPFGENKEVEREKRMNRQMGIKGVAWGFPT 326

OY 337 GAFGLMYSNYGLGFWMGSRFLVD--GAVDVGDILITVMAILIGSFGSNVSPNAQFTNA 395

DB 327 GYMWCLIFFCYALAFWYGRSLVLDGEYFPTGLIQLIFCYIAAMNIGMNSCSEIFSTG 386

OY 396 VAAAKRIFGTLIDRSPIDPNSNEKTLIDHEGHELRANKHIYTSREVTVMESVLSMP 455

DB 387 CSAASSIFQITDRQPVNDCXSGDGYKIDRKIGEIEFNVTFFHPSREPVILNLSVIK 446

OY 456 AGKTTALVPGSGSGKSVVGVVERFVMPVGRGYVLDHDKDLRLRLOOISVSGOEPV 515

DB 447 PGETTAVVSSGAGKSTALQIQFVPCRCMWTLDHDIRSLNRLRQDIGIVEDEPV 506

OY 516 LFGTTIKNIRHGLIGTRYENESDKVRELLENAAKANAHDFTALPEGETVWVGORG 575

DB 507 LFTSTIENIRLG-----REATMEDIVQAKANANVFIMALPQGFDTLVGGGG 557

OY 576 LLSGGQRIALAAVSDRKILLDLDTATSLDTSKGVQALERAEGRTTIVIAHRL 635

DB 558 QMSGGQRIALAAVIRKRIKILLDLDTATSLDSESAKVGALNKRIOHGTITISVAHRL 617

```

QY 636 STIKTANHIVVNGKIAEGTDELDEGCAVRYLY--EAQRINREKADALEDAED 693
DB 618 STYRSADVYIGFEHGRAVENGTHEELLERKGVFMVLTLOSSEDNTHKET----GIKMG 673
QY 694 LTNADI-----AKIKTAS--SASSDLGKPTTIDRTGTHKGV-----SSA 731
DB 674 TTBGDPTEPTEFSGSYODSLARASIRORSQSGLHLSHEPPL--AIGHKKSIEDKRDNDV 731
QY 732 ILSKRPPETTPKYSMTLTKLFVASENRPETPYMLIGLVESVLGGQPTQAVLYANAIST 791
DB 732 LVEEVPAP-----VRIIKYNISEMPYILVGLCAINGAVTPYLSLSQILKT 782
QY 792 LSPESQYSLRHADAFWS--LMFVYGIQFTOSTNGAFAVCSRLRRASAFRT 849
DB 783 FSLVDKQOR-----SEIYSCLFEVILGCVSLTQPLQGNFKSGGLTKRKRTGFK 838
QY 850 ILRQDIAPFDKENSSTGALTSFLSTETKLSGVSVTLGITLMTSTLGAIIITALLAIGW 909
DB 839 MLRQDIGHFEDDLKNNPGVLTTRLATDASQVQATGSGQVGMVNSFINIYAVLALFLFW 898
QY 910 KIALVCISVYPVLLACGFRFYMLAQFQSRKSLAYGSANFCEATSSITVASTLRERD 969
DB 899 KSLVLSVFPFLALSGAVQTKMLTGFASQDKELERKAGQITNEALSNIRTVAGIGVEGR 958
QY 970 VMEIYHAQIDAQGRSTLISVLRSLLYASSQALVFCVALGFWYGGTILGHHEYDIFRRP 1029
DB 959 FIKAFVEVELEKSKYKIRKANYGLCYAFSOGISFLANSAYRYGGLYEDLNTSYF 1018
QY 1030 VCFSEILFQAQSGATVFSAPDGMKAKMAAEFRRLDRKPOIDNNESEGEKLEVEGET 1089
DB 1019 RVVSSITAMATATAGRTFSTPYAKAKISARFQLDRKPPIDVYSGAEKMDNQK 1078
QY 1090 EFRNVRFRPTREPVYLAGDLTVKPGQVALYVPSGGCKSTTALLERFDALAGSLI 1149
DB 1079 DFLDKFTYPSRDQVNLGSLVSDPGQTLFVSSGGCKSTIQLLRFYDPDQGTW 1138
QY 1150 VDGKDISKINISYRSFLSVSQEPLVYOGTITKENILGIVEDVDEEFLIKACKNDANY 1209
DB 1139 IGHGSKATVNGFLRNSINIGVSEPLFDCSITMDNKIKYDNTKEISVERAIAAKAQH 1198
QY 1210 DITMISLPEGFNTVSGKMSGGQKQVAVARALLDRPKILLDLBATSALDSESEKVO 1269
DB 1199 DEVMSLPEKFTYETVNGIQSGQLSRGEKQRIARAIYRDKILLDLBATSALDSESEKVO 1258
QY 1270 ALDDAARGRTITVAHRTSTOKADVITYVPOQKIVEGTSSELYOKKGRYELY 1325
DB 1259 LALDARABGRCTIVAHRLSTIIONSDIIAMVSGVVIETGTHKILMDOKGAYIKIV 1314

RESULT 14
AB11_RABIT STANDARD: PRT: 1321 AA.
AC Q9NOV3.
DB 16-OCT-2001 (Rel. 40, Created)
DB 16-OCT-2001 (Rel. 40, Last sequence update)
DB 16-OCT-2001 (Rel. 40, Last annotation update)
DE Bile salt export pump (ATP-binding cassette, sub-family B, member 11)
DE (Sister of P-glycoprotein).
CN ABCB11 OR BSEP OR SPGP.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
CX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Liver;
RA Balasubramanian N.V., Suchy F.J., Ananthanarayanan M.;
RT "Molecular cloning and characterization of rabbit liver bile salt
RT export pump (Bsep/sggp).";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
CC INTO THE CANALICULUS OF HEPATOCYTES.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

```

```

CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
CC SITU.
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
CC ATP BINDING CASSETTE (ABC) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF249879; AAF65552.1; -
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR003439; ABC_transporter.
DR InterPro: IPR001140; ABCtransporter.
DR Pfam: PF00005; ABC_tran. 2.
DR Pfam: PF00664; ABC_membrane. 2.
DR Prodom: PD000006; ABC_transporter. 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transmembrane; Transport.
FT DOMAIN 1 62
FT TRANSMEM 63 83
FT DOMAIN 84 147
FT TRANSMEM 148 168
FT DOMAIN 169 215
FT TRANSMEM 216 236
FT DOMAIN 237 240
FT TRANSMEM 241 261
FT DOMAIN 262 319
FT TRANSMEM 320 340
FT DOMAIN 341 353
FT TRANSMEM 354 374
FT DOMAIN 375 755
FT TRANSMEM 756 776
FT DOMAIN 777 794
FT TRANSMEM 795 815
FT DOMAIN 816 869
FT TRANSMEM 870 890
FT TRANSMEM 891 911
FT DOMAIN 912 979
FT TRANSMEM 980 1000
FT DOMAIN 1001 1011
FT TRANSMEM 1012 1032
FT DOMAIN 1033 1321
FT NP_BIND 455 462
FT NP_BIND 1113 1120
FT CARBOHYD 109 109
FT CARBOHYD 116 116
FT CARBOHYD 122 122
FT CARBOHYD 125 125
SQ SEQUENCE 1321 AA; 146376 MW; 146376 CD; 6071 AA; 2 CRC64;

Query Match 33.3%; Score 2235; DB 1; Length 1321;
Best Local Similarity 38.4%; Pred. No. 4.8e-124;
Matches 504; Conservative 238; Mismatches 472; Indels 100; Gaps 18;

QY 90 KTOLEIRK-----VNISFGLMYRATKMDILINIVISTICA-----IAASTFORIMLYQISY 141
DB 31 KSRLODKKSDSVIRIGFQPLPFSSWTDIWMKMSGLCAHIGIAPGV---LILFGTWT 87
QY 142 DEFYD-----ELTKNVL-----YFVY 157
DB 88 DVFIDYDELQELKIPKACVNTIWINSSLNQVNTGTGRCGLDISEMIRFAGYAG 147
QY 158 IGIGFVIVYVSVGFITTGGAATOKIRIRYIESILRONIGFEDKLAGEVTRITADTN 217

```

Db 148 IGAIVLTGTYICFNGIAAHOIQKMRKSYFRIMMGIGWUCNSVGIKNTPEVDEN 207
 QY 218 LIODGISEKVGTLTALATEVTAIYAYVKYKIALICSSITVALVLMMGGSOFI---- 273
 Db 208 KINDSADQALFIOGTSPILFGLVGEFSQMMKLTIV-----LISVPLIGAAIIGLSV 263
 QY 274 IKTSKSLDSYGAGVYAEVYISSIRNATAFGTOCLKAKOYEAHLDEAKMGKTNQIYMG 333
 Db 264 SKETDELKAYAKAGSVADDEVYISSMRTVAAFGEKEKEREKMLVAORIGIKGIYMG 323
 QY 334 FMIGAMGLVSNYGLGFMGMSRF-LYDGAVDVGDILTVMAILIGSEFVNSPNOAF 392
 Db 324 FFGYMCCLIFCYALAFMGSKVLVEEGEYSPALVQIFLSTVIGALNIGNASPCLEAF 383
 QY 393 TNAVAAMAKIFGTIDROSPIDPYNSNEKTLDEHGHIELRNKHIYSREPVYMEVSL 452
 Db 384 AARAAASSIFETIDRPIIDCKSEDOYKLERIKGEIEFNHVFHYHSREPVKILNLSM 443
 QY 453 SMPAGKTALVPSGSGKSTVGVGVFVMPVPGCTVLLDGHDKDLNLRMLROQISLVSQ 512
 Db 444 VIRGEWTLVPGSGAKSTALDLIRFYGTGEMVTVESHDIRSSHIOWRNQGIIYEQ 503
 QY 513 EPLVFTGTYIKNRHGLIGIKYENESDKYRELINAKAMAHDETTALPEGYEVNGQ 572
 Db 504 EPLVFTHTTAETIRYQ-----REDATMEDLIQAKEMANINFTMDLPQOFTLVGE 554
 QY 573 RGFLLSGGKORIAIARAVYSDPKILLDEATSLADTSEGVQVQALERAEGRTTIVIA 632
 Db 555 GGGOMSGGKORAVIARALIRNKRKILLDMATSLADNESEMVOEALSKOHHGITYVA 614
 QY 633 HRLSTIKTANIVLVNGKTAOGTDELVDRGAYRKLV--EAQR--INEQREADALE 687
 Db 615 HRPATITADVILIGCEGAVERGTEEELEKRGVYFALVLOSRNQDGEENEKDATE 674
 QY 688 D-----ADMEDLNADIKIKITASSASSDLDGKPTI-DRTGHHKSSVSSAILSKRP 737
 Db 675 DDIPKTFSSGNTODSIRASIRO-RSKQSLSYTAHEPPMAVEDKSTHEDRK---DIDL 730
 QY 738 P-ETTPKYSILWTLKLFVASFNRPEIYMLIGLVSYLAGGGQPTOAVLYAKAISTSLP 795
 Db 731 PAQEDIEPASVVRIMK---LNAPEWPMYLLGSKGAIVNGAVPLVAFLEFSQILGTSLP 786
 QY 796 ESQYSKIRHADRWSLFVAVGIITFTOSTNGAARVCSRLIRARASTAFRTILRODI 855
 Db 787 DKE--EORSQINGICLIFVTLGCVSFTFOGTYFAKSGELLKRLKFEFRAMLQODI 844
 QY 856 AFDPKENSUGALTSLETSTKHLSSGVTGTLTMTSTLGAIIITAIIGKIALVC 915
 Db 845 GWFEDLRNSFGALTRLATDASQVQATGSOIGMNVNSFTVAVMAITALEFSMKTGTGI 904
 QY 916 ISVVPVLLAGCFYRYMIAFOFSRKSCLAYEGSANFACETSSIRTVASLTERDVMETIYH 975
 Db 905 VCFEPFLALGALQTKMLTGPASRDQALEKAGIITSEALSNITVAGIGERKFIETFE 964
 QY 976 AQDADGRTLISVLRSSLVYASSQALVFPCVALGFWYGGTLIGHHEYDIFFRVCESEI 1035
 Db 965 AELEKPKYKAIKKNANYGLGFEFGQCTIFLANSASYRGYLLINEGLHFSYFVRISAV 1024
 QY 1036 LFGNOSAGTVFSPADMGKAKNMAAEFRRLFDRPOIDNMESEKLETVGELEFRVNH 1095
 Db 1025 VLSATFALGRASSYTPYAKAKHISARFOLLDRPPIINVYSSAEKKMDFGKIDFVDC 1084
 QY 1096 FRYTPREOPVLRGLDITVARGOVALVPGSGCKSTTIALLEFFYDAIAGSIIVDGDID 1155
 Db 1085 FTYSRBDIYQVLANLSVMSRQTLAVVSGSGCKSTSIQLEFFYDPDHCKVMDIGDS 1144
 QY 1156 SKLINSYRSFLVSEOPVLYOGTICKENILGIVEDVPEEFLIKACKDANIYDFIWSL 1215
 Db 1145 RKNVIOQLRSNIGIVSOEPLVFIACISIKDNIKYGDNTOEIPMERITIAAKKQVHDFWSL 1204
 QY 1216 PEGNNTVYSGKMLSGGOKORAVIARALRDPKILLDEATSLADSESEKVVQAAIDAA 1275
 Db 1205 PEKETVNVSGSOLSHGKORIAIARAVADPKILLDEATSLADSESEKTVQVADLKA 1264

QY 1276 ARGFTTAAHRLSTIQADVIYVFDQKIVESGTHSELVOKRGYYELVNIQS 1329
 Db 1265 REGTCIYIAHRLSTIQNSDIIVAMSGOMVTEKGTHEELMAYOKGAYKLVTTGS 1318

RESULT 15
 ABIL_RAT ID ABIL_RAT STANDARD; PRT; 1321 AA.
 AC 070127;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bile salt export pump (Abp-binding cassette, sub-family B, member 11)
 DE (sister of p-glycoprotein).
 GN ABCB11 OR BSEP OR SPGP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=8212048; Pubmed=9545351;
 RA Gerloff T., Stieger B., Hagenbuch B., Madon J., Landmann L., Roth J.,
 RA Hofmann A.F., Meier P.J.;
 RT "The sister of p-glycoprotein represents the canalicular bile salt
 RT export pump of mammalian liver".
 RL J. Biol. Chem. 273:10046-10050(1998).
 CC -1- FUNCTION: INVOLVED IN THE ATP-DEPENDENT SECRETION OF BILE SALTS
 CC INTO THE CANALICULUS OF HEPATOCYTES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY, IF NOT EXCLUSIVELY IN
 CC THE LIVER, WHERE IT WAS FURTHER LOCALIZED TO THE CANALICULAR
 CC MICROVILLI AND TO SUBCANALICULAR VESICLES OF THE HEPATOCYTES BY IN
 CC STU.
 CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,
 CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN
 CC ATP BINDING CASSETTE (ABC) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MDR SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U69487; AAC40084.1; -;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR001140; ABC_tranpflrm.
 DR Pfam; PF00005; ABC_tran; 2.
 DR Pfam; PF00664; ABC_membrane; 2.
 DR Prodom; PD000006; ABC_transport; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW ATP-binding; Transmembrane; Transport.
 KW DOMAIN 1 62
 FT DOMAIN TRANSSEM 1
 FT TRANSSEM 63
 FT TRANSSEM 83
 FT TRANSSEM 84 147
 FT TRANSSEM 148 168
 FT TRANSSEM 169 215
 FT TRANSSEM 216 236
 FT TRANSSEM 237 240
 FT TRANSSEM 241 261
 FT TRANSSEM 262 319
 FT TRANSSEM 320 340
 FT TRANSSEM 341 353
 FT TRANSSEM 354 374
 FT TRANSSEM 375 755
 FT TRANSSEM 756 776
 FT TRANSSEM 776

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2003, 15:57:13 ; Search time 57 Seconds
(without alignments)
4822.225 Million cell updates/sec

Title: US-09-758-828-2

Perfect score: 6721
Sequence: 1 MSPLSTNPLSPETAMKEPAE.....YOKGRYELVNLQSGKH 1334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6704	99.7	1348	3	09Y8G1 emericella
2	5407.5	80.5	1349	3	043121 aspergillus
3	4575	68.1	1331	3	09HG55 trichophyto
4	4167.5	62.0	1353	3	08X1E2 venturia in
5	3992	59.4	1307	3	043138 aspergillus
6	2682.5	39.9	1408	3	043140 cryptococcus
7	2519	37.5	1288	13	093437 gallus galli
8	2501.5	37.2	1287	13	091586 xenopus lae
9	2496.5	37.1	1272	11	09UK64 ratius norv
10	2482	36.9	1280	6	046605 canis fami
11	2481.5	36.9	1285	6	002793 ovis arie
12	2435.5	36.2	1169	11	060502 cricetus
13	2417.5	36.0	1275	11	08R427 ratius norv
14	2273	33.8	1348	13	090235 raja erinac
15	2272.5	33.8	1163	6	09TS02 felis silve

17	2272.5	33.8	1407	5	08T9W5 dictyosteli
18	2243.5	33.4	1293	3	09Y748 emericella
19	2232.5	33.2	1321	11	088331 ratius norv
20	2226	33.1	1284	3	09Y8G2 emericella
21	2224.5	33.1	1285	10	09LX1 oryza sativ
22	2222.5	33.1	1432	5	08T9W4 dictyosteli
23	2210.5	32.9	1289	10	0941H6 coplis japo
24	2205.5	32.8	1248	10	09FHF1 arabidopsis
25	2203.5	32.8	1278	10	09FWX7 arabidopsis
26	2200.5	32.7	1292	3	09C163 rhizomucor
27	2181.5	32.5	1286	10	080725 arabidopsis
28	2173.5	32.3	1275	5	061301 haemochus
29	2172	32.3	1262	10	08RVT7 triticum ae
30	2166.5	32.2	1292	10	09M1O9 arabidopsis
31	2154.5	32.1	1313	5	09V6Z6 drosophila
32	2153	32.0	1265	5	001495 caenorhabd
33	2148.5	32.0	1287	10	08S035 oryza sativ
34	2146	31.9	1230	10	09STI3 arabidopsis
35	2141	31.9	1313	3	08X1E3 venturia in
36	2135	31.8	1294	5	045721 caenorhabd
37	2132	31.7	1229	10	09FWX8 arabidopsis
38	2128	31.7	1229	10	049749 arabidopsis
39	2128	31.7	1229	10	09STI2 arabidopsis
40	2118.5	31.5	1240	10	09LHD1 arabidopsis
41	2106.5	31.3	1323	10	09MOM2 arabidopsis
42	2105	31.3	1302	5	024851 entamoeba h
43	2098.5	31.2	1310	5	024852 entamoeba h
44	2091.5	31.1	1283	5	024393 drosophila
45	2087	31.1	1252	10	09LX0 arabidopsis

ALIGNMENTS

RESULT 1
09Y8G1 PRELIMINARY; PRT; 1348 AA.
ID 09Y8G1
AC 09Y8G1;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Multidrug resistance protein MDR (ABC-transporter).
-GN ATFD OR ABCD.
-OS Emericella nidulans (Aspergillus nidulans).
-CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
-OX NCBI_TaxID=5072;
-RN [1]
-RP SEQUENCE FROM N.A.
-RC STRAIN=W6-096;
-RA Andrade A.;
-RT "Two novel ABC transporters from Emericella nidulans.";
-RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
-RN [2]
-RP SEQUENCE FROM N.A.
-RA Nascimento A.M., Terenzi M.F., Goldman M.H.S., Goldman G.H.;
-RT "Molecular characterization of ABC-transporter encoding genes in Aspergillus nidulans.";
-RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
-CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
-DR EMBL; AF071411; AAD43626.1; -
-DR EMBL; AF173826; AAF29805.1; -
-DR HSSP; P13569; INBD.
-DR InterPro; IPR003593; AAA_ATPase.
-DR InterPro; IPR001140; ABCtransportTM.
-DR InterPro; IPR003439; ABC_transportr.
-DR Pfam; PF00664; ABC_membrane; 2.
-DR Pfam; PF00005; ABC_tran; 2.
-DR Prodom; PD000006; ABC_transportr; 2.
-DR SMART; SM00382; AAA_2
-DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
-DR ATP-binding; Transport.
-SQ SEQUENCE 1348 AA; 147467 MW; 7B0506AB631D0218 CRC64;

Query Match 99.7%; Score 6704; DB 3; Length 1348;
 Best Local Similarity 99.0%; Pred. No. 0;
 Matches 1334; Conservative 0; Mismatches 0; Indels 14; Gaps 1;

QY 1 MSPLNPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRP 60
 Db 1 MSPLNPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRP 60
 QY 61 KSSSSNNAVSNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 120
 Db 61 KSSSSNNAVSNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 120
 QY 121 TICATAA-----ASTFORIMLYOISYDEFDLTKNVLFEYVYLGIGEFVY 166
 Db 121 TICATAAAGALPLFTLLFGLSLASTFORIMLYOISYDEFDLTKNVLFEYVYLGIGEFVY 180
 QY 167 YSTVGEITGEGHATOKIREYVLESTLRONIGVEDKLAGEVTRTRTADPNLQDGISEK 226
 Db 181 YSTVGEITGEGHATOKIREYVLESTLRONIGVEDKLAGEVTRTRTADPNLQDGISEK 240
 QY 227 VGLTLTALATFYAFLIAYVYKWKALICSSSTVALVLPWGSSQFLIKYSKSLDSYGA 286
 Db 241 VGLTLTALATFYAFLIAYVYKWKALICSSSTVALVLPWGSSQFLIKYSKSLDSYGA 300
 QY 287 GGTVAEEVYSSIRNATAFGTQDKLAKOYEYHLDPAEKWGTKNQYNGFMIGAMGLMYSN 346
 Db 301 GGTVAEEVYSSIRNATAFGTQDKLAKOYEYHLDPAEKWGTKNQYNGFMIGAMGLMYSN 360
 QY 347 YGIGFPMGSRFLVDGAVDGDILTVLMAITLIGSFSLGANSPPNQAOTNVAANAALFEGT 406
 Db 361 YGIGFPMGSRFLVDGAVDGDILTVLMAITLIGSFSLGANSPPNQAOTNVAANAALFEGT 420
 QY 407 DRQSPDPYSSNEKTLDEHGEHLELNKHIYSPREVTVMEDVSLMPAGKTALVGS 466
 Db 421 DRQSPDPYSSNEKTLDEHGEHLELNKHIYSPREVTVMEDVSLMPAGKTALVGS 480
 QY 467 GSGKSTVGLVERFYMPVPGTYLLDGHDKDLNRLRQOISLVGSEPYLFGTTIYKNIR 526
 Db 481 GSGKSTVGLVERFYMPVPGTYLLDGHDKDLNRLRQOISLVGSEPYLFGTTIYKNIR 540
 QY 527 HGLIGKYENESBDKRELIENAKANAHDFETALPEGETVNGRGFLSGGQORLA 586
 Db 541 HGLIGKYENESBDKRELIENAKANAHDFETALPEGETVNGRGFLSGGQORLA 600
 QY 587 IAAVVSDEPKILLDEATSDLTDKSEGVYQAALERAEGRTTIVIAHRSTITANIVY 646
 Db 601 IAAVVSDEPKILLDEATSDLTDKSEGVYQAALERAEGRTTIVIAHRSTITANIVY 660
 QY 647 LVNGKTAEGTHDELVDRCGAYKLVQAORINEQKADALADADLTINADIAKIKTAS 706
 Db 661 LVNGKTAEGTHDELVDRCGAYKLVQAORINEQKADALADADLTINADIAKIKTAS 720
 QY 707 SASDDGKPTTIDRTGTHKSVSALISRPPTTKYSIMTLKRVASNREIETIYML 766
 Db 721 SASDDGKPTTIDRTGTHKSVSALISRPPTTKYSIMTLKRVASNREIETIYML 780
 QY 767 GLVFSVLAGGQPTQAVLYAKAISTLSPESQYSKLRHADWLSLAFVVGIIQFTQST 826
 Db 781 GLVFSVLAGGQPTQAVLYAKAISTLSPESQYSKLRHADWLSLAFVVGIIQFTQST 840
 QY 827 NGAAPVCSERLIRARSTAFRTILRODIAFEDKEENSTGALTSPLSTETKHLGVSQV 886
 Db 841 NGAAPVCSERLIRARSTAFRTILRODIAFEDKEENSTGALTSPLSTETKHLGVSQV 900
 QY 887 LCTIIMTSTTLGAIIIAIAIGKALVLCISVYPVLLAGGFYFYLAQFQSKIAYES 946
 Db 901 LCTIIMTSTTLGAIIIAIAIGKALVLCISVYPVLLAGGFYFYLAQFQSKIAYES 960
 QY 947 SANFACEATSIPTVASLRENDVWEIYHAODAOGRSTLISVRSLSLYASQALVFC 1006
 Db 961 SANFACEATSIPTVASLRENDVWEIYHAODAOGRSTLISVRSLSLYASQALVFC 1020

QY 1007 VALGFWYGGTLLGHHEYDIFRFVFCSEILFGAQSAGTVFSPADMGKAKNAAEFRRLF 1066
 Db 1021 VALGFWYGGTLLGHHEYDIFRFVFCSEILFGAQSAGTVFSPADMGKAKNAAEFRRLF 1080
 QY 1067 DRKPOIDNNSEEGEKLETVGCELEFRNVHRYPTREOPPLRGDLTVPRGOVVALVGS 1126
 Db 1081 DRKPOIDNNSEEGEKLETVGCELEFRNVHRYPTREOPPLRGDLTVPRGOVVALVGS 1140
 QY 1127 GCGSTTIALLEFPYALGSLIVDGKDISKLNINSRSLISVSEPTLYOCTIRENLI 1186
 Db 1141 GCGSTTIALLEFPYALGSLIVDGKDISKLNINSRSLISVSEPTLYOCTIRENLI 1200
 QY 1187 LGIVEDVDPEEFLIKACKDANIYDFIMSLPEGNTVYVSGKGLSGGQORVAIARALLR 1246
 Db 1201 LGIVEDVDPEEFLIKACKDANIYDFIMSLPEGNTVYVSGKGLSGGQORVAIARALLR 1260
 QY 1247 DPKILLDEATSDLSESESEVVOALDAARGTTTAAVHRSTIOKADVIYVEDGKITV 1306
 Db 1261 DPKILLDEATSDLSESESEVVOALDAARGTTTAAVHRSTIOKADVIYVEDGKITV 1320
 QY 1307 ESGTHSELVQKGRYVELVNLQSLGKH 1334
 Db 1321 ESGTHSELVQKGRYVELVNLQSLGKH 1348

RESULT 2
 043121 PRELIMINARY; PRT; 1349 AA.
 AC 043121;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Multidrug resistance protein 1.
 GN MDR1.
 OS Aspergillus fumigatus (Sartorya fumigata).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes.
 OC Eucoriales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5085;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=10AF/86/10;
 RX MEDLINE=98038972; PubMed=9373135;
 RA Tobin M.B., Peery R.B., Skatrud P.L.;
 RT "Genes encoding multiple drug resistance-like proteins in Aspergillus fumigatus and Aspergillus flavus.";
 RL Gene 200:11-23(1997).
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: U62934; AAB8658.1; -;
 DR EMBL: U62933; AAB8657.1; -;
 DR HSSP: P13569; INBD.
 DR InterPro: IPR003593; AAA_Arpase.
 DR InterPro: IPR001140; ABCtransport.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_tran; 1.
 DR ProDom: PD000006; ABC_transport; 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Transport.
 SQ SEQUENCE 1349 AA; 147723 MW; B90A86FD62C4165 CRC64;

Query Match 80.5%; Score 5407.5; DB 3; Length 1349;
 Best Local Similarity 79.1%; Pred. No. 2; 1e-310;
 Matches 1069; Conservative 114; Mismatches 141; Indels 27; Gaps 6;

QY 5 EYNPPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRKSS 64
 Db 5 EYNPPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRKSS 64
 QY 5 EYNPPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRKSS 64
 Db 5 EYNPPLSPERAMEPEPTSTTEEOASTPHADEKKILSDSAPSSSTATPADKEHRKSS 64
 QY 65 SNAASVNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 124
 Db 65 SNAASVNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 124
 QY 65 SNAASVNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 124
 Db 65 SNAASVNEVDALIAHLPEDEROYLKTOLKEIKVINSFPGLRATKMDILIMVIS 124

QY	125	IAA-----	ASFOFMTMJOISYDEFOYELRKNVLYFYVJGISGEFVNVYST	170
Db	119	IAAGALPLPFIITLFGSLASAPOGISLCTMYPHEFYHKLTKNVLYFYVJGISIAEFVNVYST	178	
QY	171	VGFITYGEMAHQKIREYLLTESILRONIGYDKDGAEBVWTRITADINLQDISEKVGIT	230	
Db	179	VGFITYGHEHLQKIRENYLEAILRQNMAYPDKDGAEBVWTRITADINLQDAISEKVGIT	238	
QY	231	LTLATITVYTAFTIAYVYKMLALICSTIYALVYLMGGSGQFIYKSKSLDYSAGAGTV	290	
Db	239	LTAATITVYTAFTIAYVYKMLALICSTIYALVYLMGGSGSRFIYKSKSIESYAGAGTV	298	
QY	291	AEEYISSRRNATAGTODKLAKOYEHVLDABEKWGRKNOYMGFMGAMFGMLSNYSGIG	350	
Db	299	AEEYISSRRNATAGTODKLAKOYETHLAEBKWGAKOYVLTGMGMGFMGFMFSYSGIG	358	
QY	351	FWMSSRFYLVGADVGVJLIVLVAAILIGSFSLGNVPNAQAFNVAANAALKIFGTRDROS	410	
Db	359	FWMSSRFYLVGADVGVJLIVLVAAILIGSFSLGNVAPNQAFTNGVAAAKIYSTITDRS	418	
QY	411	PLDPSYDNBCKTLDFECHIELRNVKHIYSPRPYVWEDVSLSPMAGKTTALVPGSGSK	470	
Db	419	PLDPSYDEGCKYLDHFEGNINIEFRNVKHITYSPRPYVWEDVSLSPMAGKTTALVPGSGSK	478	
QY	471	STVVGIVBERFMYPRGVJVLDDGHDIKDNLRMTRQOISLYSGEPVLFCTTITYKTRNRCIT	530	
Db	479	STVVGIVBERFLLPVGGOVJLDDGHDIQTLNLRMTROISLYSGEPVLFCTTIRNRCHEGL	538	
QY	531	GTRYENESDEVRLEJINAKMANAHDFETALPEGETVNVGORGFLTSGSGOKORIALARA	590	
Db	539	GTRYENESKOKIRLELVENAMNANAHDFETALPEGETVNVGORGFLTSGSGOKORIALARA	598	
QY	551	VVSDPKILLDDEATSAIDYKSEGVYOALERRAEGRTTIVIAHRSTIKTAHNIVLVNG	650	
Db	559	IYVSPKILLDDEATSAIDYKSEGVYOALDKAEGRTTIVIAHRSTIKTAHNIVAMVG	658	
QY	651	KIAROGHDELVDNGGAVRYLVAORINBOKEADAE---DADAELTNAADIAKIKTASS	707	
Db	659	KIAROGHDELVDNKGKYLYLVAORINBEKEALEADADMADEFGOEGVYTRITAVS	718	
QY	708	ASSDLDC---KPRPTDRCTGTHKSVSSALISKAPPEPTTYPKYSIMTLTKVASFNNPEIY	763	
Db	719	SSNSLDANVEKARLEMKRRTQKSVSSAVLSKVPQOFERYSLMTLYVKITGAFNRPBEGY	778	
QY	764	MLGLFVSVLAGGGOPTQAVLYAKAISTJSLPESQYSKLRHADFWSLMFEVVGIIOTIT	823	
Db	779	MLGLTFESFLAGGGOPTQAVLYAKAISTJSLPESMHKLKHADFNWSLMFEVYVIAQFIS	838	
QY	824	QSTNGAFVAVCSEKLIRASTAFRITLRQDLAFEDKRENSGALTSPJSTETKHLGSV	883	
Db	839	LSTNGTAFALCSEKLIRASTAFRSLTRQDLAFEDFRERNSTGALTSPJSTETKHLGSV	898	
QY	884	GVTJGTILMTSTYTGAMIIITIALAIGKMLAVCSVVPVLLAGCFYFYMLAOFQSRSKIA	943	
Db	899	GVTJGTILMTSTYTGAMIIITIALAIGKMLAVCSVPIILLACGFLFYFYMLAOFQSRSSA	958	
QY	944	YEGSAPACEATSSIRTVASLIRERDWEIYTHAQDLAOGRTSLIYLRSSLLYASSQALY	1003	
Db	959	YEGSASVACEATSAIRTVASLIRERDQWGYHYDQLOKQGRKSLIYLRSSLLYASSQALY	1018	
QY	1004	FFCVVALGFVYGGITLGHHEVDIRFVYVCSSEILFGAQSGAGTVSFAPADMGAKKNAAEFR	1066	
Db	1019	FFCVVALGFVYGGITLGHHEYSIRFVYVCSSEILFGAQSGAGTVSFAPADMGAKKNAAPFK	1076	
QY	1064	RLFPRKRPIDMWSBEGRKLETVGECEIEFRVWHRXYRTPREOPVLRGLDITVYRPGYVALY	1122	
Db	1079	KLFBSKRPIDMWSBEGRKLESMBECEIEFRVWHRXYRTPREOPVLRKLNLSVAPRGOTIYALY	1138	
QY	1124	GPSGCGSTTIALLEREYDAIAGSLVYDKDISKLINYSRPSLSYVSOEPLTYOGTIKE	1183	
Db	1139	GPSGCGSTTIALLEREYDADLAGVYDQKDTIKLWNVSYSRPSLSYVSOEPLTYOGTIKE	1199	
QY	1184	NILGLYEDVYBEEFLKACKDANIYDFINSLDEGNTYVGSKGMLSGGOKORVALARA	1244	

ID	NAME	PROT	1331 AA
Db 1199	NILLEGKADVDSEETLIKVCNDANTIDFWSM;PEEGDIYVSGSGWLSGGQKORVAIAIA		1258
QY 1244	LIRDRKILLIDPATSAIDSESEKVVQALDAARAGRTTAAVHRSLSTICKAPVIYVDPDG		1303
Db 1259	LIRDRKILLIDPATSAIDSESEKVVQALDAARAGRTTAAVHRSLSTIONADIIYVDPDG		1318
QY 1304	KIVSGTSHSELVOKKGRYYELVNLQSLGKH		1334
Db 1319	KIVSGTSHSELVOKKGRYYELVNLQSLGKTH		1349
RESULT 3			
ID	Q9HGT5	PRELIMINARY;	
AC	Q9HGT5		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DI	01-OCT-2001 (TREMBLrel. 18, Last sequence update)		
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
GN	Multidrug resistance protein MDR.		
OS	ATRD.		
OC	Trichophyton rubrum.		
CC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;		
CC	Oryziales; Arthrodermataceae; mitosporic Arthrodermataceae;		
OC	Trichophyton.		
RN	NCBI_TaxID=5551;		
RA	SEQUENCE FROM N.A.		
RA	Fachin A.L., Ferreira-Nozawa M.S., Maccheroni W. Jr.,		
RT	Martinez-Rossi N.M.;		
RT	"Identification and Characterization of atrd gene, a new multidrug		
RT	ATP-Binding Cassette Transporter of Trichophyton rubrum.";		
CC	Submitted (Aug-2001) to the EMBL/Genbank/DBJ databases.		
RL	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL; AF291822; AAG01549.3; -		
DR	InterPro; IPR003593; AAA_ATPase.		
DR	InterPro; IPR002106; AATRNA_LigaseIT.		
DR	InterPro; IPR001140; AbctransprtTm.		
DR	InterPro; IPR003439; ABC_transportr.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00605; ABC_tran; 2.		
DR	ProDom; PD000006; ABC_transportr; 2.		
DR	SMART; SM00382; AAA_2.		
DR	PROSITE; PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
DR	ATP-binding; Transport.		
SO	SEQUENCE 1331 AA; 145283 MW; E35B7ACE304360 CRC64;		
Query Match 68.1%; Score 4575; DB 3; Length 1331;			
Best local similarity 67.9%; Pred. No. 3e-261;			
Matches 901; Conservative 183; Mismatches 223; Indels 20; Gaps			
QY 26	EQASPTPHAADEKKIISLDLSAPSSITTATPADKREHRK--SSSSNNAVSNEVDALIAHPE		83
Db 3	EVSSEKPNQDDSDVSQKQENNPASSSSSTSDKKEKVAKKGNSDATKSTPDDIDAOIAHPE		62
QY 84	DEROVLKLTQDEIKINISIFGLMRYATKMDILIMYSTICATAIA-----AS		129
Db 63	HERELLKQQLFIPDKAKATYGTLEFRATRNDMIFLAIVASIAAGAAALPLFVYLGSLAG		122
QY 130	TFQRLMLQVSIYSDEFYDELTKNVLVYVYIGIEFVYVYVSTVGFTYTGHAQOKIREYLL		189
Db 123	TFRDIALAHRTYDEPNISILRNSLFFVYVYIGIAQIFILLYVSTVGFTYVGHITQIKRATKL		182
QY 190	ESILNONGYEPKLDAGEVTRITADTNLIQGISKVGSLITLALATFPTATIAVKKV		249
Db 183	HALIKNONGIEFKLDAGEVTRITADTNLIQGISKVGSLITLALSTFESAFIIGVRYV		242
QY 250	KALICSSITVAVLTMGGSOFIKIKYSKSDSYGAGGTVAEEVYSSIRNATFAGTODK		309
Db 243	KALICSSITVAVLTMGGISRFYVKSGMTLVSTVIGEGGTVAEEVYSSIRNATFAGTODK		302
QY 310	IAKQVEVHLDEAEKKGTQKNOIVGMIGAMFGIAMSNTGLGFWMGSRFLVDGAVDGLT		369

```
DB 303 LARQYEVHLKEARKGRRIOMLGLTIFGSMALMTSYNGLGFWMGRFLVGGTLDLSATV 362
QY 370 TVLMALILSFGSLGNVSPNAQAFPTNAVAATAKIFGTIDRQSPDLPYSNECKTLDHEGHT 429
DB 363 NILALIVISFSGIGNVAPNTQAFASASAGAKIFSTIDRVASADPDSDECDTIEENEGTI 422
QY 430 ELBNVNHITPSREYVYVMEVSLSMRAGKTTALVPSGSGSKTVYGLVEFYMPVKTUL 409
DB 423 EFGFHNHITPSREYVYVMEVSLSMRAGKTTALVPSGSGSKTVYGLVEFYMPVKTUL 402
QY 490 LQGHDIKIDNLKRWLRQOISLVSOEPLVFGTTIKNIRHGLIGTKYENESDQKRELIENA 549
DB 483 LQGRDITKTLNKLRLQOISLVSOEPLVFGTTIKNIRHGLIGTKYENESDQKRELIENA 542
QY 550 AKMANAHDEFTALPEGYETNVGGRGLSGGOKORIAIARAVYSDPKITLLDEATSLDPT 609
DB 543 AKMANAHDEFTALPEGYETNVGGRGLSGGOKORIAIARAVYSDPKITLLDEATSLDPT 602
QY 610 KSEGVQAALERAABERTTIVIAHRLSTIKTANIVYLVNGKTAEGTHDELVDGCAVR 669
DB 603 KSEGVQAALERAABERTTIVIAHRLSTIKTANIVYLVNGKTAEGTHDELVDGCAVR 662
QY 670 KLYEAOIRNEOKADALDADAE--DLTNADIAKIKITASASDDLGRKPTTIDRTGTHK 726
DB 663 QLYEAOIRNEOKADALDADAE--DLTNADIAKIKITASASDDLGRKPTTIDRTGTHK 722
QY 727 SVSSALILS-KRPETTPKYSMLTWLLEKVASFNRPETPYMLGLVFSVLAGGQPTQAVLY 785
DB 723 SVSSALILS-KRPETTPKYSMLTWLLEKVASFNRPETPYMLGLVFSVLAGGQPTQAVLY 782
QY 786 AKAISSLIPESQYSLKRLHDADQMSLMFYVGIQTITOSTNAAARVCEKILIRARST 845
DB 783 AKAISSLIPESQYSLKRLHDADQMSLMFYVGIQTITOSTNAAARVCEKILIRARST 842
QY 846 AERTILRODIAFEDKEENSTGALTSFSTETKHLSGVSGTGLTITMTSTGLAIIIAL 905
DB 843 AERTILRODIAFEDKEENSTGALTSFSTETKHLSGVSGTGLTITMTSTGLAIIIAL 902
QY 906 AIGWKIALVCISSVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSSIRTVASLT 965
DB 903 AIGWKIALVCISSVYVLLACGYRYTMAOFQSRKSLAYEGSANFACEATSSIRTVASLT 962
QY 966 RERDVEIYHAQIDNAGRTSLISVLAASSILLYASSQALVFCVALGFWYGGTLLGHHEDY 1025
DB 963 RERDVEIYHAQIDNAGRTSLISVLAASSILLYASSQALVFCVALGFWYGGTLLGHHEDY 1022
QY 1026 FREFVCFSEILFQAQAGTVEFAPDMGRKAKNAAEFRRLDFDRKPOIDMMSGEKLEFY 1085
DB 1023 FREFVCFSEILFQAQAGTVEFAPDMGRKAKNAAEFRRLDFDRKPOIDMMSGEKLEFY 1082
QY 1086 EGEIERFNVHFRYPTRPEOPVLKGLDLYVKGQYVALVGPBGCGKSTTALLERFYDA 1145
DB 1083 EGEIERFNVHFRYPTRPEOPVLKGLDLYVKGQYVALVGPBGCGKSTTALLERFYDA 1142
QY 1146 GSLVNDKXIKININISYSEFSLVSQEPITVQGTIKENILIGIVDDVPEEFLIACAD 1205
DB 1143 GSLVNDKXIKININISYSEFSLVSQEPITVQGTIKENILIGIVDDVPEEFLIACAD 1202
QY 1206 ANIYDFIMSLPEGFNTVYVSGKGLSGGOKORAVIARALLRDKPILLDLEATSLDSE 1265
DB 1203 ANIYDFIMSLPEGFNTVYVSGKGLSGGOKORAVIARALLRDKPILLDLEATSLDSE 1262
QY 1266 KYVQAALDAAGRTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQKKGRYELV 1325
DB 1263 KYVQAALDAAGRTTIAVAHRLSTIQKADVIYVFDQKIVESGTHSELVQKKGRYELV 1322
QY 1326 NQOSLGR 1332
DB 1323 HMOSLER 1329
RESULT 4
```

```
08XIF2
ID 08XIF2 PRELIMINARY; PRT; 1353 AA.
AC 08XIF2;
DT 01-MAR-2002 (TREMBLrel, 20, Created)
DT 01-MAR-2002 (TREMBLrel, 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel, 21, Last annotation update)
DE ATP-binding cassette transporter ABC4.
OS Venturia inaequalis.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes; Chaetothyriomycetes; Incertae sedis; Venturiaceae;
OC Venturia.
OX NCBI_TaxID=5025;
RN [1]
RP SEQUENCE FROM N.A.
RA Schnabel G., Schnabel E.L., Jones A.L.;
RT "Complete sequences of four ABC transporters from Venturia
inaequalis."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF375879; AAL57243.1; -.
DR InterPro; IPR003593; AAA_Atpase.
DR InterPro; IPR001140; ABCTransportTM.
DR InterPro; IPR003439; ABC_transportr.
DR Pfam; PF00664; ABC_transportr. 2.
DR Pfam; PF00005; ABC_tran. 2.
DR Prodom; PD000006; ABC_transportr; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; UNKNOWN_2.
DR ATP-binding.
SQ SEQUENCE 1353 AA; 149136 MW; 7C8E90B3D723D2C CRC64;

Query Match 62.0%; Score 4167.5; DB 3; Length 1353;
Best Local Similarity 61.0%; Pred. No.3,5e-237;
Matches 834; Conservative 184; Mismatches 290; Indels 45; Gaps 6;

QY 5 ETRNPLSPETAMRPAETSTFEQASTPHADEKKIISLDSAPSTYATPARDKERRKSS 64
DB 19 KTRGSGSETEIKDKNARQMGAPY-----GMDTPSGGKLKMKRL 64
QY 65 SNNASV-NEVDALIAHLPDERQVLKTOLEIKVNIISFGLRWYATKMDILTMVISTIC 123
DB 65 DSNVIEVKQNDPFRHLDEHEQELIRQTFIPDVKGVFTLYRASRMDMAWMWLSVFC 124
QY 124 AIAA-----ASTFORIMLYOISDEFDELTKVLYIVYIGIEFTY 165
DB 125 STVSGAMPIMTVFGLGLGFADYRKVN---ITYKOFSELSHPVLYFLYLAIGTFY 180
QY 166 VYVSTVGFITYGHNATOKIREYLYESILRONIGYPRKLAGEVYTRITADTNLIQDIS 225
DB 181 TYIMTVGFITYGRCGKIRERLTKMLKMLNIAFPDKLAGEVYTRITADTNLIQDIS 240
QY 226 KVGTLTALATFYATIAVVKYWKLLALICSTIVVALVLTMGGSQFTIKYKSSLSY 285
DB 241 KFGILTALATFISAVVIAFIKYYKLLTLITLTSVPFATITVMGVSSFFVWVRSQTE 300
QY 286 AGGTVAEEVYISIRNATATGTOCKLAKOYEVHDEAEKMGTKQIYMGFMIGMFLM 345
DB 301 KGGTIAEEVYISVRNAIAFPTQDILAKYVDSYLVIAETMGKIKOMISGAMISMTL 360
QY 346 NYGLGFMWGRFLVDGADVVDGILITVIALILISFSLGNVSPNAQAFPTNAVAATAKIF 405
DB 361 NYGLSFMWGRFLVDGDMYVSYVLYLFAVMIGAFSLGNVAPRFAKFTSVNAAQGLIF 420
QY 406 IDROSPLDPSNEGTLDHEGHIELRNKHIYPSREYVYVMEVSLSMRAGKTTALVGP 465
DB 421 IDRTSPMDPSPDGKYLEKMSGPIELRNKHIYPSREYVYVMEVYVNGVDLILPAGK 480
QY 466 SSGKSTVYGVYVERFMPARGVLYLDGHDIKDNLKRWLRQOISLVSOEPLVFGTTIK 525
DB 481 FSGKSPVYGLVERFEPYEPGGEMLDGHDIRLNLKRWLRQOISLVSOEPLVFGTTIK 540
QY 526 RHGLIGTKYENESDQKRELIENAAKMANAHDEFTALPEGYETNVGGRGLSGGOKOR 585
DB 541 RFGILGTEFEKVDPERQDILLEGAMNANHDEFTALPEGYETNVGGRGLSGGOKOR 600
```


Db	769	LEYPRLRDANFNCIMFIMIGITVSVLYSGTLEFAVSSSEKVMRAQSARFVILHIDIS	828
Qy	857	FEDEKEENTGALSTSLSTETFHLSGVSVYLTGLTILMTSTLGAALITAIAGKTLAVCI	916
Db	829	FEEDQENTTGLATLNLSTAGTELTGISGVTLTGLTIIVSVNLVASTLGAALYGMKTLAVCI	888
Qy	917	SVVPELTLACGFREFYMLAQFOSRSKLAIEGSAFNACEATSSIRTVASTITREDDVMELYHA	976
Db	869	SAVPALLMCGVVRWVWMLERFQRRAKKVAQESASSACEAASLRIVVELTITMETALQSTQA	948
Qy	977	QLDQAGTSLTSLVLRSSLLTIVASSALVFFCVAGLFWTGTGLLGHHEYDIRFEVCFSEIL	103
Db	949	QLRRQLKDDIDPIRYKSSLLTIVASSALPEFCMALGFWMYGGSLIGHGESELPQFYVCFSEVI	100
Qy	1037	FGAAGAGVFSEFADPMGAKAAAEFRRLRPRKQDIDNWSEBGRKLTVEBSELEFRVHF	109
Db	1009	FGAAGAGVFSEFADPMGAKAAAEFRKLFESDTHMASRS-KGVPAVSMKGLVEFRVVF	106
Qy	1097	RYPTREPOVPLRGDLVTPKQOYVALVPSGCGSKSTTALLERPYDAIAGSLVDKDIS	115
Db	1068	RYPSRLQPIRLRLNLTKPQFALVAVGASSGSKSTTALLERPYDPLKGVVYDGNIT	112
Qy	1157	KLTINSTRSELSVSOEPTLYQGCTIKENILGLIVEDDVEBELKACKDANIYDFINSLP	121
Db	1128	TLEMSYSRSHALLISOEPTLTQGTIRENILLGSSNTPHYDDEFLKACKDANIYDFILSLP	118
Qy	1217	EGFNTVYVSKCGMISGGCKQKQVATARALTRPKILLDDEATSSALDSSEKVVQAALDAA	127
Db	1188	QGFNTYVSKCGMISGGCKQKQKRIATARALTRPKILLDDEATSSALDSSEKVVQAALDAA	124
Qy	1277	RGRTTIAVAHNLSTIQRADVLYVFDQGIYVESGTHSELVQKRGYELVNLDS	1329
Db	1248	RGRTTIAVAHNLSTIQRADLTYLVLDQGVVBSGTHRELKRKGYYELVNLQN	1300
RESULT 6			
ID	043140	PRELIMINARY;	PRT; 1408 AA.
AC	043140:		
DT	01-JUN-1998 (TREMBLrel. 06, Created)		
DT	01-JUN-1998 (TREMBLrel. 06, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	Multidrug resistance protein 1.		
GN	MDR1.		
OS	Cryptococcus neoformans (Filobasidiella neoformans).		
OC	Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;		
OC	Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.		
OX	NCBI_TaxID=5207;		
FN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-M1-106;		
RX	MEDLINE=98072425; PubMed=9409767;		
RA	Thornewell S.J., Peery R.B., Skatard P.L.;		
RT	"Cloning and characterization of CEMDL: a Cryptococcus neoformans		
RT	gene encoding a protein related to multidrug resistance proteins.";		
RL	Gene 201:21-29(1997).		
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL; U62930: AAC48890.1; -;		
DR	EMBL; U62929; AAC4889.1; -;		
DR	InterPro: IPR003593; AAA_ATPase.		
DR	InterPro: IPR002106; AATRNA_ligaseII.		
DR	InterPro: IPR001140; ABCtransprtTM.		
DR	InterPro: IPR003439; ABC_transportr.		
DR	Pfam; PF00664; ABC_membrane; 2.		
DR	Pfam; PF00005; ABC_tran; 2.		
DR	ProDom; PD000006; ABC_transportr; 2.		
DR	SMART; SM00382; AAA; 2.		
DR	PROSITE; PS00339; AA_TRNA_LIGASE_II.2; UNKNOWN_1.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 2.		
KW	ATP-binding; Transport.		
QC	SEQUENCE 1408 AA; 152143 MW; CEBDA34A09A6F69F8 CRC64;		

Query Match	39.9%	Score 2682.5	DB 3	Length 1408
Best Local Similarity	43.9%	Pred. No. 1.9e-149		
Matches 597	Conservative 233	Mismatches 449	Indels 81	Gaps 16
QY	33	AADEKKILSDLSAPSSTTA--TPADKEHRKRS--SSNNAVSNEYDALLAHLPEDERQVL	89	
Db	69	AAEGKISRSIAASDTLNSPLEK---PISNAFSKSHYKKSKEFDLKRKKKEEEK	125	
QY	90	KQOLE---IKNISPFGIMRYATKMDILINVISITCAIAAS-----T	130	
Db	126	NKEKREASVLPPVSFFALFRAPALEIITAMVLGLVAAMASCOPLMLILGRLLTSTP	185	
QY	131	FORIMYOISY-----DEYDELTKVLYFVYLIGEPTVVSSTGEFY	175	
Db	186	NAVAIANQISOGGLIPERSALQAADKDKTOSGHNAITLMAIGIMFLATVYFIMV	245	
QY	176	TGEHATQKIREYYSLELRONIGYFDKIGAGEVYTRITADIMLIDGISSEKGLTETLATA	235	
Db	246	TEGLMSKRIRREYRLAAVLNQETAYEDDLGAEVANRIQTDCHLVQEGISEKALVPQYAG	305	
QY	236	TVYVAFIAYVYKMLALICSTIVALVITMGCGSOFIITYSKSLDSYAGGVAEEVI	295	
Db	306	TFVCGFVLAFVSPRLAGALVSLPVMICGIMMTAMAKYGTALADHAKAGSLAEVI	365	
QY	296	SSIRNATFNGODKLAKQYEVHLDEEKKGSTKNQIYMGFMIGAMGLMSTNGLEFGWMS	355	
Db	366	GSIRYVOAGKKKILGDKFADHIEOSKIYGRKGSIPEFGSLIMFVYAAALAFYYCG	425	
QY	356	RLVYDGAVDVDILTVLMAILIGSFLGVSNNQAFTVAVAAKIFCTIDROSPLDY	415	
Db	426	ILVSGQADSGIVINVFMSILGTSMSMLAEELAAVYTKARAALKFATIDRVAILDSA	485	
QY	416	SNEGTLDHFEGHIELRNVKHIYPSRPEVTVMEDVSLSPAGKTTALVGPSSGSKSTYWG	475	
Db	486	SEEGKRPGLGRGEISFENYKFFHPSPSPILPKGFTTPEAKTALVAGSSGSKSTYVS	545	
QY	476	IYERYMPRGVYLLDDHDKIDLNRWLROQISLYSQEVLFGTTIYKIRIGLIGTYE	535	
Db	546	ILERYDPVSGVAKIDGRIRSLNINWLRQOGLVLSQEPTEFTTVGRNGVEGLIGSRE	605	
QY	536	NSESEKVRLEIENAKMANADFTALPEGYETYNVGOREFLLSGGOKRIATARAVSPP	595	
Db	606	NASLEKEFELVYKACVDANNAHFINKLPQGYVTMGGERMLLSSGGOKRYALARAIVSDP	665	
QY	556	KILLIDEATSALDYSKSEGVYQALERRAEGRTTVIAHRLSTIKTAHNTVLYNGKIAEQ	655	
Db	666	RLILDEATSALDPTSEGIYODALDKASGRFTTIAHRLSTIRADRIYVGGVELEQ	725	
QY	656	GTHDELV--DRGAYKRIYEAQRINQOKENDALE--DADAEDLTNADIATIKTASASSDD	713	
Db	726	GSHNDLLENENPYQVILNNOKLAEAAEAALQVDDIDEDPDA-----VFIGGSPMQ	779	
QY	714	GKPTTIDRTGTHKSYSSA---ILSKRPPEY-----PKYSILMTLL-----KFVASF	756	
Db	760	EEDKOLHRAVYGRSLASTIAMDDIOAKRAEYVAGEKITSFGLYRLRLMNSADKFT---	836	
QY	757	NKPEIPYMLIGLVFSYLAGGGQPTQAVLYYAKAISTLSPEOSYKILRHDAFWMSLMEFY	816	
Db	837	-----YIAFIAIACAGMYVPSLAILIFGKRLSOFEIQDP--ALRLHLSLSALMYFIT	887	
QY	817	GIQPTTOSTNGAFAVCSERLIRARSTAFETILKODIAFFDEKENSALTSLSTSTET	876	
Db	888	ALAAAFVIEFOSAGSRAGMWDNGVLRKKLPATLRHDIEMWDERBNSGAVTSMIADQP	947	
QY	877	KHLGVSQVLTGTLIMTSTTGAATITIALAGWKALVACISVVPVILACGFRTFMADQF	936	
Db	948	QAVQGLFEPRLGTIVQASCATLIGCGCTIGCYGPLLALIGIACIPILVSGGYIRLVVYAK	1007	
QY	937	QSRSLKLEAGSNAFCEATSSIRYVASTLRREDVWEIYHAOLDAGGRSLVLSLSSLLY	996	
Db	1008	DQRMKMLAAASHLASAAGAKVYTAASLTREDVYKRIYSEALKAPMKLNFRSISQCLF	1067	
QY	997	ASSQALVEFVYALGFWYGGTLIGHHEYDIFRFYVCFSELFGAOGAGTVSFAPDMGAK	1056	

Dh	1068	AASGGLFECIALVFYIGALMITIDAKYSTAFSTVINSTSYERSIQAGNFTFPDASKAN	1127
Qy	1057	NAAAEFRRLFDKRPQIDNMSESECKLE--TWSEIEIFRNVHYFPTRRBQPYLRGIDLTV	1119
Dh	1128	SSAASFIRSTIDNEPAIJAENSEGKVLDHKHVGHVRLEGVHFYPRPGCVRLRNLTIDY	1181
Qy	1115	KPQGYVALVSPSGCGKSTTALLEREPDALIASILVDKDISLNTNSRFLSLVTSQBP	1177
Dh	1188	PAGTYVALVPBSGGKSTQTOMERFYDFPLAGVTLLDGIDIKELNASTYSQSLSQBP	1241
Qy	1175	TLVYGITKENILGIVE--DDVEEFLIRACKDNATYDFMTSPDEGFNTVSGSKGMLSG	1233
Dh	1248	TLVAGTIRPNILLGANKPIEEVYQDELDAACKRANITYDFVSLPDGFTEVGKGSQLSG	1307
Qy	1233	GOKORAVAIARALLDPKILLIDENTSALDSESEKVVQAALDAAAKRGTIAVAHRLSTIO	1299
Dh	1308	GOKORIAIARLIARNPVLLILEDTATSALDSQSEKKVQEAOLDRAKAAGRRTIAIHRLSIO	1363
Qy	1293	KADIVYVFDGKIYESTGSTHELVQKKGRYYELYVNDSLGK	1332
Dh	1368	HSDRIYFSEGRVAEHGTHELLAKKGYYELVQMOLSR	1407
<p style="text-align:center;">RESULT 7</p>			
ID	093437	PRELIMINARY;	PRT; 1288 AA.
AC	093437;		
DC	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DE	ABC transporter protein.		
GN	CMDBI.		
OS	Gallus gallus (Chicken).		
OC	Eucharyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Ancorosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
CC	Gallus.		
RX	NCBI_TaxID=9031;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=INTESTINE;		
RX	MEDLINE=99209805; Pubmed=10195430;		
RA	Edelmann H.M.L., Duchek P., Rosenthal F.E., Foeger N., Glackin C.,		
RA	Kane S.E., Kuchler K.;		
RT	"Cndrt, a chicken p-glycoprotein, confers multidrug resistance and		
RT	interacts with Estradiol.";		
RL	Biol. Chem. 380:231-241(1999).		
CC	-1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.		
DR	EMBL, AJ009799; CA08835.1; ..		
DR	HSSP; P13569; IMBD.		
DR	InterPro: IPR003593; AAA_ATPase.		
DR	InterPro: IPR001140; ABCtransportTM.		
DR	InterPro: IPR003439; ABC_transportr.		
DR	InterPro: IPR000504; RNA_rec.mot.		
DR	Pfam; PF00664; ABC_membrane; 4.		
DR	Pfam; PF00005; ABC_tran; 4.		
DR	PRODom; PD000006; ABC_transporter; 2.		
DR	SMART; SM00382; AAA; 4.		
DR	PROSITE; PS00211; ABC_TRANSPORTER; 4.		
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_2.		
KW	ATP-binding; Transport.		
FT	CHAIN	2	1288
SC	SEQUENCE	1288 AA;	141917 MW; CB258A5F2826DBC CRC64;
<p>Query Match 37.5%; Score 2519; DB 13: Length 1288;</p>			
<p>Best Local Similarity 41.2%; Pred. No. 7.2e+140;</p>			
<p>Matches 554; Conservative 230; Mismatches 451; Indels 110; Gaps</p>			
Qy	32	HADEKILISLAPSSITTATPADKEHPRSSSSNNNAVSVNEYEDALTAHLPEDEROVLKT	91
Dh	2	HSEDK-----QRHTVNGNYEAINSNQDD-----PEDKKGKKK	37
Qy	92	OLEEKVNISSFGLMRATKMDIILNIVISTICALAAASTFORIMLY-----	137

[illegible]


```

Db 1079 SSGCGKSTVVOQLERFYDPLSGELVFPDDIDAKTINIQWMLASHIGIVSOEPLTFPTAEN 1138
Qy 1185 ILGLYEDVDYPEEFLLIACKDANIYDFIMSLPEGEFTVSGSGKMLSGGOKORAVIARAL 1244
Db 1139 IAGDGNREYSHEEIIISNAASISHSFIDSLPEKYNTRVGDGTOLSGGOKORAIARAL 1198
Qy 1245 LRDPKILLDEATSAIDSESEKVVQALDAAAGRTTIAVAHRSTIOKADVIYVPEOQK 1304
Db 1199 IRKPQIILLDEATSAIDSESEKVVQALDAAAGRTTIAVAHRSTIOKADVIYVPEOQK 1258
Qy 1305 IVESGTHSELVQKRGYELVNLQS 1329
Db 1259 VIEGTHQQLLAKGFGYSLVNVQS 1283

RESULT 8
ID 091586 PRELIMINARY; PRT: 1287 AA.
AC 091586;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, last sequence update)
DT 01-JUN-2002 (TEMBLrel. 21, last annotation update)
DE Multidrug resistance protein.
GN XEMDR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95322451; PubMed=7599185;
RA Castillo G., Shen H.J., Horwitz S.B.;
RT "A homologue of the mammalian multidrug resistance gene (mdr) is
RT functionally expressed in the intestine of Xenopus laevis."
RL Biochim. Biophys. Acta 1262:113-123(1995).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: U17608; AAA75000.1; -
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR001140; ABC_transporter.
DR InterPro: IPR003439; ABC_transporter.
DR Pfam: PF00664; ABC_membrane_2.
DR Pfam: PF00005; ABC_tran_2.
DR Prodom: PD000006; ABC_transporter_2.
DR SMART: SM00382; AAA_2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW SEQUENCE 1287 AA; 141505 MW; 06E95073C5771415 CRC64;
SQ

Query Match 37.2%; Score 2501.5; DB 13; Length 1287;
Best Local Similarity 42.6%; Pred. No. 7.8e-139;
Matches 545; Conservative 215; Mismatches 455; Indels 63; Gaps 12;

```

```

Qy 366 GDILITVMAILIGSEFSIGNSPNAQFTNAVAARAKITFTIDROSPLDPYSNEGTLDHF 425
Db 339 GSYLTFEFAVIGAFAGQSPINIEAFANRGAAYITFINIIDNPQIDSEKGLPKDKI 398
Qy 426 EGHLELNVAHYSPREVMEDVSLMAGKTALAVGSGSKSVVGLVEFYVPR 485
Db 399 KGDIEFNVAIFYPSRKDIQVLLKGLNLIPSGKVALVSGGCKSTVYDLOEFYDPE 458
Qy 486 GTVLLDHDIDKIDLRLNRQOISLVSOEPLYFTTIRKIRHGLIGTYENESDEKREL 545
Db 459 GVTLDQDQIRSLNIRLREIIGVSOEPLTFDTIDNIRYQ-----REDYKKE 509
Qy 546 IENAKAANAHDITLALPEGETVNGQRFLLSGGOKORAIARAVYSDPKILLDEATS 605
Db 510 IERATKEANAYDFIMKLPDLLETILVERGQTLSSGGOKORAIARALVRNPKILLDEATS 569
Qy 606 ALDPTKSEGVQALERAERERTTIVAHRLSTIKTANIVLVNKAIEGDTDELVDG 665
Db 570 ALDPTSEAVVQALDKRERRTTIVAHRLSTIRNANAIGFPGVTEGSHRELMERG 639
Qy 666 GAYRKLEAQRINEQKEADALDADAE-----DLTNAIDAKIKTASASSDDGKRP 719
Db 630 GYFNLVTLQFVETSKOTE--EDLETHIYEKKIPYTHHNLVRKSSRNTITSKYPETE 687
Qy 720 DRGTGHSVSSAILSKRPPTTPKYSLSMTLLKFAVSENPBEIPYMLGLVFSYLAGGQP 779
Db 688 D-----KEVDEE--EKKEEGPPVSPFFKVM-----LNKREPFYFVGVICAMINGAT 726
Qy 780 TOAVLYAKAISTSLPESQSKLRHNDWFSLMEFYVGIITQTSINGAFAVCSRLI 839
Db 737 AFAITFSRIIGVAPGVSQ--MSESSMYSLFLAIGVSPFTFLQGTFFKAGILIT 793
Qy 840 RRASTAFRTILRODIAFEPEENSTGALTSFSTERKHLGVSYGLITILMTSTTGA 899
Db 794 MRRLISGFKMLQGEIGWFDKSNSTGALTLRLATDSOVQAGTRRLAALANVANTGT 853
Qy 900 AIIIALAIGKALAVCISVPLVLAGCFYFYMLAOFQSSSKLAVESANFACEARISIR 959
Db 854 AIIISFYGWQTLTLIAIYPIVAAAGLVEMKMFAGHAKDKRELEKAGKISTDAVLN 913
Qy 960 TVASLTPREPWELYNHOLDAGRTSLISVRSLSLYASSQALVFCVNLGFWYGTLL 1018
Db 914 TVVSLTRERKFEAMTEKSLGCPYRNSIKKALHGLTYGLSOAHNVLCICWFVSLAAYLV 973
Qy 1019 --GHHEYDIERFVCESEILGASQASGYVSPADMGKAKNAAEFRLEFRKPDQIDMS 1076
Db 974 VEGILMKLD--EVFLVSAIYVIGAMALGQTSFAPDITKAMISAHIFSLIERYPQIDYS 1031
Qy 1077 EBEKLETVEGELEFRNVHFRYPTREPQVLRGLDITVRFQYVALVGPSCGCKSTTIAL 1136
Db 1032 DQGEKPRKNCNGNVYFQVNFNYPTRPDITVLOGLDIDISVKGETILAVYSSGCGKSTTVSL 1091
Qy 1137 LEFFYDAIGSLVDGDKSLNINSYRSLISVSOEPLTYOCTIKENTILGLVEDDYPE 1196
Db 1092 LEFFYDPEBEVLDGLSVANLNIQVRAQMGIVSOEPLTFDQSIDNIAVGNMKNKRVQ 1151
Qy 1197 EFLIKACKDANIYDFIMSLPEGEFTVSGSGKMLSGGOKORAVIARALDRPKILLDEA 1256
Db 1152 EEIETAKAKENHISFLESLLDKNTRYGDKGTOLSSGGOKORAIARALIRKPKILLDEA 1211
Qy 1257 TSAIDSESEKVVQALDAAAGRTTIAVAHRSTIOKADVIYVPEOQKIVESGTHSELVQ 1316
Db 1212 TSAIDSESEKVVQALDAAAGRTTIAVAHRSTIOKADVIYVPEOQKIVESGTHSELVQ 1271
Qy 1317 KGRGYELVNLQSLGKH 1334
Db 1272 LKGVYFSLVTLQ--LGH 1286

RESULT 9
ID 09JRK64 PRELIMINARY; PRT: 1272 AA.
AC 09JRK64;

```


DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Multidrug resistance protein 1a.
 GN PGY1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR; TISSUE=LIVER;
 RA Hootveld G.J.E.J., Wilms J.W.J., Hagenbuch B., Jansen P.L.M.,
 RA Meijer D.K.F., Muller M.;
 RT "Cloning and functional characterization of the rat multidrug
 RT resistance protein Mr1a."
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL: AF257746; AAF69007.1;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR001140; ABCtransportTM.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00664; ABC_membrane. 2.
 DR Pfam: PF00005; ABC_tran. 2.
 DR ProDom: PD000006; ABC_transport. 2.
 DR SMART: SM00382; AAA; 2.
 DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
 DR ATP-binding; Transport.
 SO SEQUENCE 1272 AA; 140326 MW; B5C1DF89DFC986C5 CRC64;

Query Match 37.1%; Score 2496.5; DB 11; Length 1272;
 Best Local Similarity 42.2%; Pred. No. 1.5e-138;
 Matches 536; Conservative 225; Mismatches 460; Indels 49; Gaps 10;

QY 90 KTOLEIKVNISEFGRLKRVATKMDILIMVISTICAIAASTFORIM----- 136
 DB 23 KKEKKKKPVAIVLTMRITGMDLREYMLGTLAIHIGALPLMLVFGDMTDFANVG 82
 QY 137 -----YQIS--YDFEYDELTKNVLYFVYLGIEFVTVVYGVFTYTGHAQKIREY 187
 DB 83 NNSMSFYNNMTDIYAKLEDEMTYAYYYTGAGAVLVAVYQVSLMCLAGROQHKIRK 142
 QY 188 YLESILKQNTIGYFDKLCAGEVYTRITADTNLIQDISERKGLTLTALTEVTAIYV 247
 DB 143 FFRHINMOEIGMEFVHVGELNTRLTLDVSKINGIGDKIGMFQAMATFEFGFLIGPTR 202
 QY 248 YKRLALICSSSTIYALVLTMGSGSOFIITKSKLSDSGAGTVAEVISSIRNATGCTQ 307
 DB 203 GMLTLTILIAISPVGLSAGIWAKLISFDFKLAAYAKAGAAEYLAALRTVIAGGQ 262
 QY 308 DKLAKQYEVHLDEAEKMGFTKNOIYMGFMIGAMFGIAMSNTGLGFMGSRFLVDAVGD 367
 DB 263 KKELEKRNNNLEEKRRIGIKKAITANISMGAFLITIASYALATWYGSIVLSKEYITGQ 322
 QY 368 ILTVLAAILIGSESLGVNSPNAQATNAVAANAIIKFTIDROSLDPLDPSYNGKTLDFHEG 427
 DB 323 VLVTFEVLIGAFESVQASPIEAFAANARGAAYEVFSIIDKPSIDPSKSGHPRDNTG 382
 QY 428 HIELRNKHIYPSREPTVMEVDSLMPACKTALVGPSSGSKTVGVGLVRFMPARGT 487
 DB 383 NLEKRNHFESYPSRKDVQILKGLMKVSGQTVLVNSGGKQRIATRLRLDEATPIGE 442
 QY 488 VILGDHDIKDLNRLMLRQOISIVSQEPVLFQTTIYKNIHMLIGTKYENESDKVRELIE 547
 DB 443 VSIDGQDIRITINRYKLEIGIYVQSEPVLPATTAENIRYC-----KENTYMD-----IE 493
 QY 548 NAAKMANAHDFITALPEGYETNGQRFLLSGGQKQRIATARAAYVSPDKILLDEATSAT 607
 DB 494 KAVKANAAYDFIMKLPKFTLVGEBRGAQLSGGQKQRIATARALVRPKILLDEATSAT 553
 QY 608 DTSESEGVVAAALRAAGRTTIYIAHRLSTIKTAHNTVLVNGKIAAGTDELYDKGA 667
 DB 554 DTSEEAVVQAAALDKARGRITTIYIAHRLSTVRNADVIAFGDGIIVBQGNDEILMRKGI 613

QY 668 YKRLVEAQRINEQEA--DALEDADAEDELINADIKTATASSASDDLDEKPTTIDRTGTH 725
 DB 614 YKRLVMTQAGNEIENELGNEACESKDGID--NVDS---SKSGSSLJRSTRKSIRSPH 668
 QY 726 KVSASAILSKRPETTPKYSIMTLTKFVASPNRPEIPMLIGLVFSLVAGGQPTQAVLY 785
 DB 669 DQDELSTREKLEADDVPPASFWRLK---LNSTEMPVVGVCAITNGILQAPASITF 724
 QY 786 AKAISTLS--LPSSQSKLRHDADFWSLMPFVVGIIQFTQSGNCAFAACSEELIRRA 842
 DB 725 SKVGVFTKNDPTETQ---RQNSMLSLFLILIGLISFTITFTFQGTTFGAGELTKRL 780
 QY 843 RSTAFRTILRODIAFFDKEENSTGALTSFSTERKHLSGSVGLTGLITMTSTLGAALI 902
 DB 781 RYVWFKSLNRDIDIMFDDPKMTGALTTLRLNDAQVAGAGSRILAVITQUNANGICII 840
 QY 903 IALIGKTLALVCISVVPVLLACGFYRYMLAQOSRSKLAYEGSANFACATSTRTVA 962
 DB 841 ISLIYQWQTLTLLALVPIIAAGVEMKMLSGQALKDKKELEGSGKIATEAIEFRTVY 900
 QY 963 SLTREBDVWEIYHQAQDAGRTSLISVRSLSLVAASQALVFCVAGLGMVGTLLGHNH 1022
 DB 901 SLTREQKFEIYTAQSLQIPYNNALKKAHVFGITTSFTQAMVFSYACFRGAYLVAREL 960
 QY 1023 YDIFREFESELIFGAQASAGTVEFAPDMGKANAAEFRFLDRKPDQIDMSEEGEKL 1082
 DB 961 MFEFENVLVFAIYFGAANAQVQVSFAFDYAKAKVSAHIIRIETKPEIDSYSEGLKP 1020
 QY 1083 ETVEGEIEFRNVHFRPTREPOVYLRGIDLTVKQGYALVSPSGCKSTTALLERFD 1142
 DB 1021 NMLGNNVAFNQMVFNPYPRMIPVYQGLSTLEVKKGQTLALVSGSGCKSTVQQLERFD 1080
 QY 1143 AIAQSILVDGDKIDKLNINSTRSLISVSOEPTLYQGIKRNILLGIVEDVPEBELKA 1202
 DB 1081 PMAGVFLDEKREIQOLANQMLRAHGIYSQEPILFDCSIANINAVGNSRVVSHETVKA 1140
 QY 1203 CKDANIYDFIMSLPEFNTVVGSGKGLMSGQKQKQVATARALLDPKILLDEATSDLS 1262
 DB 1141 AKENINHFIDSLPEKYNTRVGDGQTOISGGQKQRIATARALVRPKILLDEATSDLT 1200
 QY 1263 ESEYVQAALDAAARGRTTIYVAHRLSTIQKADVIYVFDGKIYESGTHSELVOKKGRY 1322
 DB 1201 ESEYVQEAALDKAREGRFCIYIAHRLSTIQNADLIYVQNGQVYEHGHQOOLAQGIYF 1260
 QY 1323 ELVNLQSLGR 1332
 DB 1261 SMVSVOAGAK 1270

RESULT 10
 Q924L6 PRELIMINARY; PRT: 1272 AA.
 AC Q924L6;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE P-glycoprotein.
 GN MDRIA.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY;
 RA Yamazaki M., Ieake B.F., Kim R.B.;
 RT "Molecular Cloning of Rat Mr1a cDNA."
 RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF286167; AAF63023.1;
 DR InterPro: IPR001140; ABCtransportTM.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00664; ABC_membrane; 2.

DR Pfam: PF00005: ABC_tran. 2.
 DR ProDom: PD00006: ABC_transporter. 2.
 DR PROSITE: PS00211: ABC_TRANSPORTER; UNKNOWN_2.
 SQ SEQUENCE 1272 AA; 140343 MW; D9A17457362DA0ED CRC64;

Query Match 37.0%; Score 2486.5; DB 11; Length 1272;
 Best Local Similarity 42.0%; Pred. No. 5.9e-138;
 Matches 536; Conservative 223; Mismatches 462; Indels 49; Gaps 10;

QY 90 KQLEIKVNIISFGIMRVATKMDILIMVISTCAIAAATFORIML----- 136
 DB 23 KKKKKKRAVSVLTFMRVAGMLDRFYMILGLTAAIIHIALPLMLVFQDMDFANWG 82
 QY 137 -----YQIS--YDFEYDELTKNVLKYFVYLIGFEYTVVSVGVITGEHATQIKREY 187
 DB 83 NNSTSFYATDIAELEMETTYAYYTGIGAGVLIIVAIQVLSLCLAGROIHKIRCK 142
 QY 188 YLESIKRONIGYFDKLGAGVTTRITADTNLIDGISEKGLTLLTALATFVTAIAYK 247
 DB 143 FPAIINQELGMDVDVHDELNTRLDVDSKINEGIDGKMFQMAATFEGFLLGETR 202
 QY 248 YKRLALICSTIYALVLTMGGSQFIYKSKSLDSYGAGTVAEEVTSIRNATAFGTQ 307
 DB 203 GMLTLVLIATISPLVGLSAGIMAKILISFDEKELQAKAGAAEVLAIATVIAFGQ 262
 QY 308 DKIAQYEVHLDPAEKMGTRKNOIVMGFMAGMLMTSNTGLCFMWSRFLVDGAVGD 367
 DB 263 KELEERYNNLEAKRIGIKKALTANISMGAFLLIYASVALAFMYGTSLVISKEYTIQ 322
 QY 368 ILTVLMAIISFSLNSQAFNNAVAANKIFGTIDROSLPDPYSEGTLDHLEG 427
 DB 323 VLTVEFVSLGAFVQASPNIEAFANAGAAVEVSIIDNKSIDSFSKSGKRPDIQ 382
 QY 428 HIELNVAHIYPSRPVYVMEVDSLSMPAGKTTALVPGSGSGKSTVGLVERFEPVGR 487
 DB 383 NIEFKNIHSYPSRKDVQILKILNLYKSSQVALVGNSSCGKSTVQLQRLYDPIEGE 442
 QY 488 VLLDGDIDKDLNRLRQOISLVSOPVLEGTIVKINIRGLIGTYENSESEKVELTE 547
 DB 443 VSLDGDIDITNRYLREITIGVSOEPLFATTIAENIRG-----RENTMDE---LE 493
 QY 548 NAKMANADFTALPEGETYVNGRGFLLSGOKORIAIARAVSDPKILLDEATSL 607
 DB 494 KAVKENAADFIMKLPKHKDVLGERGADLSGOKORIAIARAVNPKILLDEATSL 553
 QY 608 DPKSGVQALERAAGRTTIVIAHRLSTKTANIVLVNKGIAEQGTHDELVDRCGA 667
 DB 554 DTESEAVVOAALDKAREGRTTIVIAHRLSTVRNADVIAGFDGVIXEQNHDELMREKI 613
 QY 668 YKRLVQAQRIINQKKA--DALEDADAEDLTNMDIAKIKTASASSSDDLGKPTIIDRTGTH 725
 DB 614 YKRLVQAQRIINQKKA--DALEDADAEDLTNMDIAKIKTASASSSDDLGKPTIIDRTGTH 725
 QY 726 KSVSSALISKRPPETPKSLMTLTKFVASFNRPPIYMLIGLVFSGGGGPOAVOY 785
 DB 669 DDDGELSTKEALDDDVPRAPSWRIK-----LNSTEPIYVGVFCALINGGLOPASITIF 724
 QY 786 AKAISTLS--LPESQYSLKRDADFWSLMFVVGIIQFTIOTSTONGAFAVCSERLIRRA 842
 DB 725 SLVVGFTKNDPEID-----RONSNFLSFLFLIGIISFTFELQGFTEGKAGEIILTKRL 780
 QY 843 RSTAFETILRODIAPEKRENSGTALSTSLSTETKHLSEVSVYTLGITLMTSTTLAAII 902
 DB 781 RIMVFESMLRQDISWDDPKNTGTALTTRLANDAAOVKATGSRILAVTIQNIAMLTGTII 840
 QY 903 IALAIQMLKALVLCISVVPVLTACGFYRFYMLAOFQSRKILAYGSANFACETSSIRTYA 962
 DB 841 ISLIGMQLTLLLAIVPIIATAGVEMKSLGQALKKKELEGSKITAEALIENTRIYV 900
 QY 963 SLTREDDWEIYHAQIDAGRTSLISVLSSLLYASSQALVFECVALGFWYGGTLLGHH 1022
 DB 901 SLTREDDWEIYHAQIDAGRTSLISVLSSLLYASSQALVFECVALGFWYGGTLLGHH 1022
 QY 901 SLTREDDWEIYHAQIDAGRTSLISVLSSLLYASSQALVFECVALGFWYGGTLLGHH 1022

QY 1023 YDIFREFVCESELLFGAQSAGTVSFAPDMGKAKMAAEFRRLFDKRPQIDNMSSEGETL 1082
 DB 961 MTEENVVLASALVFGAMAGVSSFPADYAKAKVASASHIRIETIKPEIDSVSTGLK 1020
 QY 1083 ETVGEELFERNVHFRTYRTPREGQPVLRGLDITVPRGVVALVYSGGCKSTTIALLEFYD 1142
 DB 1021 NMLEGANKFQVMEVNPTRNIPVLOGLSLEVKGGTALVYSGGCKSTVOLLLEFYD 1080
 QY 1143 AIAISILVDDKIDSKLININSRFLSVSOEPLLYOGTIENTILLGIVEDVDEEFLIK 1202
 DB 1081 PMAGTVFLODKELKQNLVQCLRAHLGIVSOEPLIFDCSIAENIATYDGNVSHHEILVKA 1140
 QY 1203 CKRANIDEFTMSLPEGEFTVYVSGGMLSGGOKORVIAIARALLRDPKILLDEATSLDS 1262
 DB 1141 AKRANHOFIDSLPEKXNTVPRDKGTQLSGGOKORIAIARALVQPHILLDEATSLDT 1200
 QY 1263 ESEKVOAALDAARGTTLAVAHRLSTOKAVIYFPODGKIVSESTHSELVOKGRY 1322
 DB 1201 ESEKVOAALDKAREGRTTIVIAHRLSTIOMADLIIVIONGVKREHGTQPLAOKGIYF 1260
 QY 1323 ELVNLQSLK 1332
 DB 1261 SMVSVQAGAK 1270

RESULT 11
 046605
 ID 046605: PRELIMINARY; PRT: 1280 AA.
 AC 046605:
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TEMBLrel. 21, Last annotation update)
 DE Multidrug resistance P-glycoprotein.
 GN MDRL.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Placentalia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Puel O., Lepage J.F., Alviernie M., Gallier P., Pigneur T.;
 RL Submitted (JAN-1998) to the EMBL/Genbank/DBJ databases.
 CC 1-SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 DR EMBL; AF045016; AAC02113.1; -
 DR InterPro: IPR003593; AAA_ArParse.
 DR InterPro: IPR001140; ABCtransportTM.
 DR InterPro: IPR003439; ABC_transport.
 DR Pfam: PF00664; ABC_membrane; 2.
 DR Pfam: PF00005; ABC_transporter; 2.
 DR ProDom: PD000006; ABC_transporter; 2.
 DR SMART; SM00382; AAA; 2.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
 KW ATP-binding; Transport.

Query Match 36.9%; Score 2482; DB 6; Length 1280;
 Best Local Similarity 42.0%; Pred. No. 1.1e-137;
 Matches 538; Conservative 223; Mismatches 458; Indels 62; Gaps 8;

QY 90 KQLEIKVNIISFGIMRVATKMDILIMVISTCAIAAATFORIML----- 137
 DB 23 KKKKKKPTVSTFAMFRYSNMILRYMVGTAIIHGAALPLMLVGNMTDFANWG 82
 QY 138 -----QISYDEYDELTKNVLKYFVYLIGFEYTVVSVGVITGE 178
 DB 83 ISNKTFTPIINESTNNTOHFINLHEEKTYAYYSGIGAGVLYAAIYQVSWCLAG 142
 QY 179 HATOKIREYLSILRONIGYFDKLGAGVTTRITADTNLIDGISEKGLTLLATFV 238
 DB 143 RQLKIRKQFPAIINQELGMDVDVHDELNTRLDVDSKINEGIDGKMFQMAATFEGFLLGETR 202
 QY 239 TAFITVAYKMLALICSTIYALVLTMGGSQFIYKSKSLDSYGAGTVAEEVTSI 298

```

Db 203 TGFVGTTPGKLTLLVLAISPVLGSAIWMKILSSFTKELLAYAKAGAAVEVLAAI 262
QY 299 RNATJAGTOOKLAKOYEVHIDEAEKMGCTKNOIYVGMIGAMFGMLSNYGIWMGSRFL 358
Db 263 RYVIAAGGOKKELEKRYKNLEAKRIGIKAITANISIGAAFLITAYSTALATFYGSLV 322
QY 359 VDGAVDVGDILTVLMAILLGFSLSGNVSPNAQFTNAAAKIFGTIDRQSPIDPYSNE 418
Db 323 LSSEYITIGVLTVEFVSLIGAFSIGQSPSEIAFANRGAAYEFKIIDKPKSIDYSKS 382
QY 419 GKILDFEGHIELRNKHAIPSPREYVMDVSLSPAGKTALVGGSGSKSVGLVE 478
Db 383 GHRPDIKGNLEKRNHFSPSRKEKILKGLMKVQSGQTVALVGNSSGCKSTTVQLMQ 442
QY 479 RFMAPRGVTLDDGHDKIDLMLRQOISLVSEOPVLFGTITKYNRHGIGTKYNES 538
Db 443 RLVDPTDGMVCIQDODRTTNVRLHREITGVSOEPLVFTTIAENIRIG-----RENYT 497
QY 539 EDRVRELLENAAKMANADEFTALPEGYETNVGGRFLSSGQKORITAAVAVSDPKIL 598
Db 498 MDE-----IEKAVKANAYDFIMKLPNKFDPVLGEBGARLSGQKORITAAVAVRNPRL 553
QY 599 LLDPATLDTKSGCVQALERAEBRTTVIAHRLSTIKTANIVYLVNGKIAEGCTH 658
Db 554 LLDPATLDTSEAVVQVALDKARKGRITVIAHRLSTVAMADVIAGFDGVIKGNH 613
QY 659 DELVDGAGVRKILEAORINEOKFADALDADADLTNADIAKIKTASASDDDGKPTT 718
Db 614 DELMKEGKIYFKLVTMOTRGNELB---LENNYGESKSESDDLKSPDSS-----L 663
QY 719 IDRTGTHKSVSALLSKRP-----PETPKYSILWTLKLFVASFNPBEIPLYMIGLVS 771
Db 664 IKRSTRSRSHAPGGOKRIGTKEDLNENVPVSFWRLK-----LNTSEMYEYVAGICRA 719
QY 772 VLAGGGPTQAVLYAKAISTLSLPESOYSKLRHADPWSLMEFYVGIIOFTQSTNGAAF 831
Db 720 IINGGLPAPSIIFSRIGITFTRDEDETK-RQNSNMFVFLVIGIISFTFPLQGTFF 778
QY 832 AVCESELRIRARSTAFRTIIRADIAFDKEENSGALTSFTEKHLISGSGVTLGITL 891
Db 779 GKAGEILTKRIRYVFSMRLQDVSFWDPRKNTTGALTTRILANDAAOYKGAIGSLAVIT 838
QY 892 MTSTTLGAIIIALAIGMKLALVCIISVAVPVLACGFREFYMLAOPQSRKSLAYEGSANA 951
Db 839 QNTANLGCTGIIISLYGMOQLTLLALVPIALAGVEMKMLSGOALKDKKELGAGRTA 898
QY 952 CEATSSIRTVASTLRQDWEIYHAQDAGRTSLISVNSLLYASSQALVFECVALGF 1011
Db 899 TEALENFRVAVSLTRQCFEYMAQSLQVPRNSLRKAHIFGVFSITQAMMYFSYACGF 958
QY 1012 WYGTTLGHHEDYDFFRFVQSEILFGNOSAGTFSFADPMGKAKNMAAERRLJEDRPQ 1071
Db 959 RFGVLYANENPMQDVLVFSALVFGMAAGVSSFPADAKKAVSAHYIMITIEKSP 1018
QY 1072 IDNMSGEKLEVEGEIEFRNVHFRYPTRDEPVLRLGLDITVRPGOVALVGPSSGCKS 1131
Db 1019 IDSISPGIKRNTLEGANTFENVEYFNPTRPDIPLQGLSLEVKKGGQALVYSSGCKS 1078
QY 1132 TTIALLEFPDIALGSLVDGKDISKLNINSYRSLVSOEPLLYOCTIKENILIGIVE 1191
Db 1079 TVVOLLERFVDPLAGSVLIDKEIKHLNVOMLRAHLGIVSOEPLIFDCSIENIAYGNS 1138
QY 1192 DDVEEELFIKACDKANIDYFMSLPEGFNVVSGKMLSGOKORVATIALADPKIL 1251
Db 1139 RYVHEELMQAKKANIHHEIETPEKYNTNVGDKGTQLSSGQKORITAAVAVQPHIL 1198
QY 1252 LLDATATLDSSEKRVQVALDAARGTTIAVHRLSTIKADVIYVFDGKIVESGTH 1311
Db 1199 LLDATATLDSSEKRVQVALDAARGRTCVIAHRLSTIQNADLIYVFOGKVENHETH 1258
QY 1312 SELVQKKGRITELVNLQSLK 1332

```

```

Db 1259 QQLIAQKGIYFMSISVQAGAK 1279
RESULT 12
ID 002793 PRELIMINARY; PRT; 1285 AA.
AC 002793;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Multidrug resistance protein-1.
GN MOR1.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID:9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Longley M., Crawford A.M.;
RT "Ovine mdrl gene."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL, U78609; AAB58489.1; -.
DR InterPro: IPR003593; AAA_Atpase.
DR InterPro: IPR001140; ABCtransport.
DR InterPro: IPR003439; ABC_transport.
DR Pfam: PF00664; ABC_membrane; 2.
DR Pfam: PF00005; ABC_tran; 2.
DR ProDom: PD000006; ABC_transport; 2.
DR SMART: SM00382; AAA; 2.
DR PROSITE: PS00211; ABC_TRANSPORTER; 2.
DR ATP-binding; Transport.
KW SEQUENCE 1285 AA; 142020 MW; 90135617C44856F CRC64;
SQ
Query Match 36.9%; Score 2481.5; DB 6; Length 1285;
Best Local Similarity 41.8%; Pred. No. 1.2e-137;
Matches 535; Conservative 230; Mismatches 467; Indels 49; Gaps 7;
QY 81 LPEDEROVL--KTQLEEKVNIISFFGLWRVATKMDIILMVSTICAI----- 125
Db 17 LKROKRRFSKKDEKKEKRPVSTFTMFRYSNMWLDRLCMVGTIAIITHGAGLPLMLVYF 76
QY 126 -----AAASTFORIMLYQISYDEYD-----ELTKNLYEVYVILGIEFYVYV 168
Db 77 GDMTDSFAGAGNLGNITLSNISTSTIDRTREYKGLKEKEMTYAYYSGIGAGVLIAYI 136
QY 169 STVGFIYGEHATQKIREYIEBSILKQNIYGFDKLAGEVTTTRITADTNLIQDGISERV 228
Db 137 QVSEFCLAGROVHRIRKQFPHAIMQDEIGWFDVHDVCKMTLRINDVSKINEGIGDKIG 196
QY 229 LTLTALFVAFITAYVYKWKIALICSTIYALVLMWGGSSQFIIKYSKSIDSGAGG 288
Db 197 MFPQAMATFFGFIIGFTTGMNLTIVYIILPVGLSAIYAKKIISFTDKKLAYAAAG 256
QY 289 TVAEVIESSIRNATPAGTQDRIAKOYEVHLDDEAEKMGCTKNOIYVGMIGAMFGMLSNY 348
Db 257 AVAEVLAIAITVYALFGQKKELEKRYKNLEAKRIGIKAITANISGAFLITAYSYA 316
QY 349 LGFMGSRFLVDGAVDGDILTVLMAILLGFSLSGNVSPNAQFTNAAAKIFGTIDR 408
Db 317 LAFWGTSLVLSREYSIQVLTVEFVSLIGAFSIGQASPNIEAFANAGAAVEYRKIDN 376
QY 409 QSPIDPYNENEGTLDHFEGHIELRNKHAIPSPREYVMDVSLSPAGKTALVGGSGSKS 468
Db 377 KPSIDISYNTGKRPKNIGNELEFRNVHFRYPSRNEVKILKGLNKLVGGSGQTVALVGN 436
QY 469 GKSTVGLVERFVMPDGTVLDDGHDKIDLMLRQOISLVSOEPLVFGTTIYKNIKRG 528
Db 437 GKSTVGLMORLYDPTBEGVNSIDGDIRTINVRVRLREITIGVYSGQEPVLEATTIAENIR 496
QY 529 LIGTYENESDEKVELLENAAKMANADEFTALPEGYETNVGGRFLSSGQKORITAA 588

```


QY 928 YRFMTLAOFOSRKLATVEGSANFACENTSSIRIVASLITREDVWEIYHAQIDAGCRISLI 987
 DB 764 VEMKMLGOALDKKREKELSGKITEALEIENFRIVSLTRECKEFMYAQSIQIPRNALK 823
 QY 988 SVLSSLLVASSQALVEFCVAGLGFVGGTGLGHHEYDIFRFVEGSESLFGAQSAGTFS 1047
 DB 824 KAHVFGITFSTQVMMTFSTACRFPGAYLVARELMPFENVLVLSALVFPGAMAAGVSS 883
 QY 1048 FAPDMKAKNAAMAEFRRLFEKRPQIDMWSEGEKLEVEGEIEFRNHFRTPTRECPVL 1107
 DB 884 FAPYAKAKVASHIIMIEKVPISIDYSTGGLKPNTEGVKNEVEYFNPTREDIYVL 943
 QY 1108 RGLDLVYKPGQYVALVGPSCGKSTTIALERFDALAGSLVLDGKDISKININSYRFL 1167
 DB 944 OGMLIEYKKGQTLALVSSGSGKSTVVOILERFYDPAAGTFLDGEVNOALVOMLRAHL 1003
 QY 1168 SLVSEPTLYOGTIEKENILGLIVEDDVEEFLIRACKRANITYDFMSLPEGFNTVYSGK 1227
 DB 1004 GIVSOEPLIFDCSTAEINATYDNGSRVVSQDEIERAKKANIHOFESLPDKYNTFVGDKG 1063
 QY 1228 GMLSGGQKORVAIRALRLDRPKILLDEATSAIDSESEKVVQALDAARGTITIAVHR 1287
 DB 1064 TQLSGGQKORAIARALVRQHIILDEATSAIDSESEKVVQALDKAREGRTCTIAVHR 1123
 QY 1288 LSTIQKADVIYVPOGKIVSGTSELVQKGRYIELVNLQSLGK 1332
 DB 1124 LSTIQADLIYVIONGKVEHGTQOQLAKGIFYSMVSQAGAK 1168

RESULT 14

ID Q8R427 PRELIMINARY; PRT; 1275 AA.

AC Q8R427; 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ATP-binding cassette protein 81b.
 OC Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
 RA Yabuuchi H., Ishikawa T.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY082609; AAL92458.1; -
 KW ATP-binding
 SQ SEQUENCE 1275 AA; 141222 MW; 700DE5F6A0E0E19D CRC64;

Query Match 36.0%; Score 2417.5; DB 11; Length 1275;
 Best Local Similarity 40.8%; Pred. No. 7e-134;
 Matches 517; Conservative 231; Mismatches 461; Indels 57; Gaps 8;

QY 94 EEIKVNISFPGIMRATKMDLILIVISTICAIAASTFORIML----- 136
 DB 26 KEKRPVIGIFGMFRYADWLKICMAIGTIAIHHGTLLPLMLVGYMTDSFTQAEIRLL 85
 QY 137 -----YQSYDEFDLKNLYLFYVLIGIFGVYVYVSGVIGTIGEHATQKI 184
 DB 86 PSVTNQSINSTQVSDSSLEEDMAMVAYTGTIGAGVLIYAVISLWCLAGAQIIRTI 145
 QY 185 REYLESILRONIGYFDKIGAGEVTRITADTNLIQDISEKVGTLTALATFAVAFIA 244
 DB 146 RQKPFHALMNGEIGFVNDAGELNTRLDVSKINDIGDKLGMFPQSITTFSGFIIG 205
 QY 245 YKTKWKLALCSSTIVALVLTWGGSQPTIKYSKSLDSYGAGGVAAEVEVISIRNATF 304
 DB 206 FLSGKTLVLVIAVPLIGLSSAMWAKVLTSTNKELOAYAKAGAVAAEVEIARIVIAF 265
 QY 305 GTQDCLAQOYEVLDEAEKMGTKNOIYMGFMIGAMFGLMYSNYGFGMWSRFIVDGAVD 364
 DB 266 GQOKELERKYNLEAKRVGIIKAITANISIGIAYLIVTASYALAFWYGTSLVLSNEYS 325

QY 365 VGDILFTVMAILIGSFLGANSVPAQAFTNAVAAAKIFGTIDRQSPIDPYSNECKTLDH 424
 DB 326 IGOVLTFEFSILGTFSIGHLAPNIEAFANAGAAVEIFKIIDNPSIDSTGKHKRPS 385
 QY 425 FEQHLELNKVIHYSREYVIMEDVSLSMRPAKGTALVGPBGSKSTVGLVEFYIPV 484
 DB 386 IMGNLEFRNVFNYSRSEVKILKGLNLKVKSGQVVAALVNGSGCKSTVQLOLQYDPI 445
 QY 485 RGTVLDDHDKDLNRLMROQISLVSOEPLYFGTITIKNIHGLIGTKYENESDQYRE 544
 DB 446 EGEVSDQODIRTIWKLIRELIGVSOEPLYFATTIAENITKG-----REMYTDE--- 497
 QY 545 LIENAKANAHDFITLALPEGYETNVGQGFLLSGGQKORAIARAVVSDPKILLDEAT 604
 DB 498 -IEKAVKRNAYDFIMKLPKHNTELVGERGAQLSGGQKORAIARALVRNPKILLDEAT 556
 QY 605 SALDTRSEGVQOALERAEGRTTIVIAHRLSTITANIVLVNGKIAEGSTHBEYDR 664
 DB 557 SALDTRSEBAVQOALDKAREGRTTIVIAHRLSTVNRNADVIAGFDGVTVEQNHBEIMKE 616
 QY 665 GGAYRKLVQAO-RINE---QKADALDEDAEDLTNNDIAKIKTASSASDLDGKPTTI 719
 DB 617 KGIYFKLVMTQIRGNEIEPGNNAYESQSDPTGASELTSEKSKPLIRSRIRSRHRODE 676
 QY 720 DRTGTHKSVSAILSKRPETTPKYSLWTLKLFVASFNNPEIPLYMIGLVFSLVAGGQP 779
 DB 677 RLUSKEDVD-----EDVPMVSFWQILK-----LNISEPPLYLVGLCAVINGQIQP 723
 QY 780 TQAVLYAKAISTLSLPEQOYSKLRHDADFWSLMEFVVGIIQITTOSTNGAARAVSERTI 839
 DB 724 VFAIYFSKIVGVFSDDDHETKQR-NCNLFSLFLVGMGISVYTFPQOFARAGEIILT 782
 QY 840 RRASTAFRTILRODIAFDFKESNGALTSLSTETKLSGVSGVTLGTLITMTSTTGA 899
 DB 783 KLRIRNFKSMLRQDISWFDHDKNTGSLTTTLASDASVNAKMSRLAVVQNVANLCT 842
 QY 900 AIIIALAIGMKLALVCISVVPVLLACGFYRFYMLAOFOSRKLATVEGSANFACENTSSIR 959
 DB 843 GIIISLVGMQTLTLVLIPIVLIGITEMKILSQALKDKKELEISGKIATEIENFR 902
 QY 960 TVASLITREDVWEIYHAQIDAGCRISLISVLSLLYASSQALVFCVAGLGFVGGTILG 1019
 DB 903 TVVSLTRQKFEFTMAQSLQIDPYRNALKRAHVGITFAQTQAMIFYSAACRFPGAYLVA 962
 QY 1020 HHEYDIFRFVCFSESLFGAQSAGTYFSPAPDMKAKNAAMAEFRRLFEKRPQIDMWSEEG 1079
 DB 963 RELMTFENVMLVFSAVVPGAMAAGNTSSPAPIYAKAKVASHIIRIKETIPEIDISTG 1022
 QY 1080 EKLETVGEIEFRNHFRTPTRECPVLKGLDLVYKPGQYVALVGPSCGKSTTIALER 1139
 DB 1023 LKPNMLEGVKNFNGVFNFPTRPNIPVLOGLSFEYVKGQTLALVSSGSGKSTVQOLLER 1082
 QY 1140 FPDALAGSLVLDGKDISKININSYRFLSVQOEPLVYOGTIEKENILGLIVEDDVEEFL 1199
 DB 1083 FYNPAGVTFDGGKELKQNLQVOMLRAHIGVSOEPLIFDCSTIENATYDNGSRVVSHEBI 1142
 QY 1200 IKACKDANITYDFMSLPEGFNTVYSGKGMGSGGQKORVAIRALRLDRPKILLDEATSA 1259
 DB 1143 VAAARBANIHOFIDISPEKYNTRVGDGTQLDSGGQKORAIARALVRQHIILDEATSA 1202
 QY 1260 LDSESEKVVQALDAARGTITIAVHRSLSTIQKADVIYVPOGKIVSGTSELVQKGR 1319
 DB 1203 LDTESEKVVQALDKAREGRTCTIAVHRSTIONNDLIYVIONGKVEHGTQOQLAAGK 1262
 QY 1320 RYVELV 1325
 DB 1263 IYFSW 1268

RESULT 15

ID Q90235 PRELIMINARY; PRT; 1348 AA.

